



HAL
open science

Evaluation ergonomique de produits et services : le cas du web marchand

Ghulaine Clouet

► **To cite this version:**

Ghulaine Clouet. Evaluation ergonomique de produits et services : le cas du web marchand. Interface homme-machine [cs.HC]. Université Paris VIII Vincennes-Saint Denis, 2005. Français. NNT : . tel-00109013

HAL Id: tel-00109013

<https://theses.hal.science/tel-00109013>

Submitted on 23 Oct 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Référence bibliographique pour citation :

Clouet, G. (2005) : *Evaluation ergonomique de produits et de services : le cas du Web marchand*. Doctorat d'ergonomie. Université Paris 8, Saint-Denis.

UNIVERSITE PARIS 8 – SAINT DENIS
U.F.R. DE PSYCHOLOGIE : PRATIQUES CLINIQUES ET SOCIALES

THESE

pour obtenir le grade de
DOCTEUR DE L'UNIVERSITE PARIS 8
EN ERGONOMIE

présentée et soutenue publiquement

le 14 décembre 2005

par

Ghulaine **CLOUET**

**EVALUATION ERGONOMIQUE DE PRODUITS ET DE SERVICES :
LE CAS DU WEB MARCHAND**

Directeur de thèse : Pierre RABARDEL

Tutrice : Viviane FOLCHER

Jury :

François DANIELLOU	Professeur, Université Bordeaux 2 (rapporteur)
Valéry NOSULENKO	Professeur, Institut de Psychologie de l'Académie des Sciences de Russie, Moscou (rapporteur)
Imad SALEH	Professeur, Université Paris 8
Viviane FOLCHER	Maître de conférence, Université Paris 8
Pierre RABARDEL	Professeur, Université Paris 8
Isabelle LAMBERT	Ergonome, France Télécom R&D

Mes remerciements vont tout d'abord à Pierre Rabardel et à Viviane Folcher pour leur suivi attentif et guidage dans ce travail, ainsi que l'autonomie qu'ils m'ont accordée tout en soutenant mes choix.

Je voudrais également remercier François Daniellou et Valéry Nosulenko qui ont accepté d'être rapporteurs, ainsi qu'Imad Saleh de sa participation dans ce jury.

Sincères remerciements à Valéry Nosulenko pour avoir accepté de se déplacer de Moscou, pour m'avoir initiée à l'approche systémique de la psychologie appliquée à la qualité perçue et pour ses diverses explications sur le terrain comme à l'Université Paris 8.

Merci aux membres du service d'évaluation de France Télécom R&D qui m'ont accueillie pour accomplir ce travail.

En particulier, merci à Isabelle Lambert pour son encadrement, sa participation à ce travail et pour avoir su trouver des solutions aux différentes difficultés.

Merci à Hélène Poirier, Marc Legrand, Nathalie Pican pour leurs conseils et guidage dans l'analyse des données.

Merci à André Poussier pour avoir su trouver des solutions techniques astucieuses.

Un très grand merci à Gilles Rouquet pour avoir animé les entretiens et les focus groups et pour son aide dans les mises au point méthodologiques.

Plusieurs professionnels de la veille-marketing et de l'entreprise marchande par Internet ont permis que des entretiens soient réalisés et m'ont accordé de leur temps. Merci en particulier à Roger Brusq pour sa réelle implication dans la construction de l'outil d'aide à l'élaboration d'évaluations.

Merci à Marco, Patrice, Gilles, Eric et Mireille pour leurs corrections de ce texte.

SOMMAIRE

2.5. L'évaluation sans l'utilisateur	37
2.5.1. Les modèles formels	38
2.5.2. Les méthodes d'inspection de l'utilisabilité	39
2.5.3. Avantages et inconvénients des méthodes d'inspection de l'utilisabilité	40
2.6. L'évaluation subjective et sensorielle	42
2.6.1. L'histoire de l'évaluation sensorielle.....	43
2.6.1.1. Des débuts relatifs à la psychophysique	43
2.6.1.2. La perception et l'activité mentale	43
2.6.2. Les méthodes d'évaluation intégrant l'élicitation de descripteurs ou d'attributs	45
2.6.2.1. Activités collectives et individuelles d'élaboration de descripteurs dans le domaine des produits alimentaires	45
2.6.2.2. Méthodes et techniques d'élicitation d'attributs en ergonomie	47
<i>La repertory grid et techniques de catégorisation</i>	47
<i>La technique SHIRA</i>	48
<i>La technique d'élicitation associative</i>	49
<i>L'approche systémique de la psychologie appliquée à l'évaluation de la qualité perçue</i>	49
2.6.3. L'évaluation à partir de dimensions déjà construites.....	53
2.7. L'évaluation par questionnaire	54
2.7.1. Les questionnaires « prêts à l'emploi ».....	54
2.7.1.1. Les questionnaires pour évaluer l'utilisabilité.....	54
2.7.1.2. Les questionnaires pour évaluer la satisfaction et/ou l'utilisabilité subjective.....	55
2.7.1.3. Prise en compte de la satisfaction et des dimensions hédoniques et émotionnelles en ergonomie	56
2.7.1.4. Variété des dimensions d'évaluation dans le domaine du Web.....	57
2.7.2. Les études exploratoires et la construction d'un questionnaire	58
2.7.2.1. Les études exploratoires	58
2.7.2.2. La construction d'un questionnaire	59
3. Les questions de recherche.....	60
 CHAPITRE 2	 63
LES DEUX VERSANTS DE L'ÉVALUATION ERGONOMIQUE : PRODUCTION DE RESULTATS ET LEUR UTILISATION.....	63
1. La question de la pertinence d'une évaluation.....	65
2. Les connaissances sur l'élaboration de données subjectives et verbales pour la construction d'un questionnaire	66
2.1. Dimensions d'évaluation produites par des utilisateurs versus par des experts .	67
2.2. Comparaison de l'entretien individuel et du <i>focus group</i> pour la construction d'un questionnaire.....	68
2.2.1. Les connaissances actuelles	69
2.2.2. Nature des contenus produits	70
2.2.2.1. Qualité ou niveau de détail des données.....	70
<i>Les caractéristiques de l'entretien individuel</i>	70
<i>Les caractéristiques du focus group</i>	71

2.2.2.2. Données spécifiques à chaque technique.....	73
2.2.2.3. Variété des différents aspects abordés et points essentiels dégagés	74
2.2.3. Difficulté d'exploitation des données	75
2.2.4. Productivité en termes de temps et de coûts dans l'une et l'autre technique	75
2.3. La question du nombre de sujets	76
2.4. Comparaison d'une situation de verbalisation avec versus sans utilisation du produit/service.....	77
3. D'une logique de production de résultats d'évaluation à une logique d'utilisation des résultats	78
3.1. Activité de conception et activité d'usage : deux axes de recherche complémentaires	78
3.2. L'approche instrumentale	81
3.3. La prise en compte de l'activité des professionnels du Web marchand dans l'évaluation.....	83
3.3.1. Des évaluations, des guides de recommandations mais peu d'études sur les « besoins » des professionnels de la transformation.....	83
3.3.2. Les questions de recherche dans ce domaine.....	84
4. Les terrains et la démarche de recherche	85
4.1. L'organisation de France Télécom et les entités concernées par la recherche ...	85
4.1.1. La branche FTA de France Télécom.....	86
4.1.2. La branche opérationnelle FT1 et son service de veille-marketing (VM)	87
4.1.3. La filiale de France Télécom, l'entreprise Sitemarchand.com (MARCH) et ses professionnels du Web marchand	88
4.2. Les activités de recherche au sein des deux terrains.....	89
4.2.1. Les activités sur le premier terrain de recherche	90
4.2.2. Les activités sur le second terrain de recherche.....	90
4.2.2.1. Des difficultés qui occasionnent une démarche « très empirique ».....	91
4.2.2.2. Récapitulatif des entretiens menés.....	91
4.2.3. Trois types de sujets.....	92
5. Conclusion.....	93
CHAPITRE 3	95
ANALYSE COMPARATIVE DE L'ENTRETIEN INDIVIDUEL ET DU <i>FOCUS GROUP</i> POUR LA CONSTRUCTION D'UN QUESTIONNAIRE D'EVALUATION.....	95
Repères pour le chapitre.....	97
1. Introduction.....	99
2. Objectifs de l'étude	100
2.1. Description du processus dans lequel s'inscrit l'étude	100
2.2. Les questions renseignées dans l'étude.....	102
3. Description de l'étude	103
3.1. Les sujets.....	103
3.2. La méthodologie	104
3.2.1. Verbalisations avec et sans utilisation de produits/services	104
3.2.2. Méthodologie de l'entretien individuel.....	105

3.2.2.1. Phase de verbalisation sans utilisation de produits/services.....	105
3.2.2.2. Phase de verbalisation avec utilisation de produits/services	106
3.2.3. Méthodologie du focus group	107
3.2.3.1. Phase de verbalisation sans utilisation de produits/services.....	108
3.2.3.2. Phase de verbalisation avec utilisation de produits/services	109
3.3. L'animateur ou intervieweur.....	111
3.4. Le dispositif des entretiens.....	111
4. Données et résultats sur la production de contenu verbal et de dimensions.....	112
4.1. Analyse des données	112
4.1.1. Elaboration d'une liste exhaustive de dimensions utilisateurs	112
4.1.2. Description de l'analyse des données et du codage des unités sémantiques	118
4.1.2.1. La variété de dimensions différentes et les idées (différentes) détaillées	118
4.1.2.2. Les idées nouvelles par rapport aux dimensions élaborées par les experts	119
4.1.2.3. Les idées spécifiques à chaque technique.....	120
4.1.2.4. Les méga-dimensions ou points essentiels générés par chaque technique	120
4.1.2.5. Les idées difficiles à interpréter.....	121
4.1.2.6. La situation de verbalisation sans versus avec utilisation du site marchand	122
4.2. Résultats	122
4.2.1. Quantité d'idées nouvelles par rapport aux dimensions élaborées par les experts	122
4.2.2. Variété des dimensions et quantité d'idées détaillées.....	123
4.2.3. Quantité d'idées spécifiques à chaque technique.....	126
4.2.4. Méga-dimensions ou points essentiels générés par chaque technique ...	126
4.2.5. Quantité d'idées difficiles à interpréter	128
4.2.6. Les performances de la situation de verbalisation sans versus avec utilisation de sites marchands	128
5. Données et résultats sur la productivité des techniques.....	129
5.1. La productivité de chaque technique	129
5.1.1. La productivité du nombre total d'idées	130
5.1.1.1. La productivité du nombre total d'idées en termes de ressources humaines	130
5.1.1.2. La productivité du nombre total d'idées en termes de coûts financiers... 134	134
5.1.2. La productivité des dimensions utilisateurs, des sous-dimensions et des détails	137
5.2. L'effet de saturation de la productivité du nombre total d'idées en entretien individuel	140
6. Synthèse des résultats.....	143
7. Performances des deux techniques utilisées en vue de la construction d'un questionnaire	145
7.1. Intérêt des données utilisateurs par rapport aux données experts.....	145
7.2. Avantages de l'entretien individuel pour l'élaboration des dimensions utilisateurs	146

7.2.1. Avantages concernant la nature et l'exploitabilité des idées produites ..	146
7.2.2. Avantages concernant la productivité.....	147
7.2.3. Recommandation sur le nombre de sujets.....	148
7.3. Avantages du <i>focus group</i> pour la construction d'un questionnaire de type « général »	148
7.3.1. Le focus group, une solution pour raccourcir le processus d'évaluation ?	149
7.3.2. Type de questionnaire produit à partir des méga-dimensions de chaque technique	151
7.4. Les limites de l'étude	152
7.4.1. Concernant la méthodologie du focus group	152
7.4.2. Niveau de généralisation possible à partir de l'étude	153
8. Intérêt des deux techniques pour la conception de produits/services	154
8.1. Combiner des entretiens avec une séquence d'utilisation individuelle et des <i>focus groups</i>	155
8.1.1. Les données pour la conception issues des entretiens individuels.....	155
8.1.2. Les données pour la conception issues des focus groups	155
8.2. Les discussions de confrontation pour étudier une variété d'usages et de motivations.....	157
9. Conclusion.....	158
 CHAPITRE 4	 161
DEUX POINTS DE VUE SUR L'ÉVALUATION.....	161
Repères pour le chapitre.....	163
1. Introduction	165
2. Présentation de la démarche suivie	166
2.1. Une démarche « non méthodologique »	166
2.2. Présentation des professionnels sollicités	167
2.2.1. Les professionnels côtoyés sur le premier terrain de recherche : les professionnels de l'évaluation	167
2.2.2. Les professionnels interviewés sur le second terrain de recherche : les professionnels de la transformation.....	167
<i>Le premier professionnel du service de veille-marketing, spécialiste du Sitemarchand.com.....</i>	<i>167</i>
<i>Le second professionnel du service de veille-marketing, non spécialiste du Sitemarchand.com.....</i>	<i>168</i>
<i>Le responsable du Webmastering du Sitemarchand.com</i>	<i>168</i>
<i>Deux ingénieurs concepteurs de sites Internet</i>	<i>168</i>
3. La construction d'une évaluation à partir des dimensions utilisateurs	169
3.1. L'artefact « dimensions utilisateurs » dans ce processus de construction	169
3.1.1. Un contexte donné	169
3.1.2. Une série de médiations pragmatiques	170
3.2. Des facteurs économiques et organisationnels.....	172
3.2.1. Les facteurs économiques	172
3.2.2. Les facteurs organisationnels.....	172

3.3. La question de la « vérité » et de l'utilité d'une évaluation.....	175
4. La compréhension des professionnels de la transformation quant à la liste de dimensions utilisateurs	176
4.1. Conduite itérative d'entretiens	176
4.1.1. Avec les deux professionnels de la veille-marketing.....	176
4.1.2. Avec les deux ingénieurs : entretiens approfondis avec le Sitemarchand.com.....	179
4.1.3. Avec le responsable du Webmastering du Sitemarchand.com : entretien de validation.....	179
4.2. La transformation de surface de la liste exhaustive de dimensions utilisateurs	180
4.2.1. Un nouveau concept d'artefact pour favoriser une lecture et une compréhension plus efficaces	181
4.2.1.1. Un concept d'artefact au format intermédiaire	181
4.2.1.2. Le regroupement des dimensions pour structurer cet artefact	182
4.2.2. Modification de l'artefact au niveau du vocabulaire	185
4.2.2.1. Les sources de confusion entre dimensions et modification des intitulés	185
4.2.2.2. Ajout d'éléments du langage professionnel.....	186
4.3. Insuffisance de la transformation de surface.....	187
4.3.1. Fournir concrètement ce sur quoi l'évaluation d'un site marchand peut porter	187
4.3.2. Préciser les objectifs de l'évaluation liés aux caractéristiques du produit/service.....	188
4.3.3. Opérationnaliser les dimensions pour constituer un outil de conception d'une évaluation.....	189
5. Deux mouvements épistémologiques distincts	191
5.1. Des situations, finalités et compétences différentes	191
5.2. Deux mouvements épistémologiques pourtant distincts.....	192
6. Conclusion.....	193
CHAPITRE 5	197
VERS LA CONCEPTION D'INSTRUMENTS POUR LES PROFESSIONNELS DU WEB MARCHAND.....	197
Repères pour le chapitre.....	199
1. Introduction	201
2. Cahier des charges de la boîte à outils	203
2.1. L'objectif de l'évaluation sous la forme d'une question	203
2.2. Découpage en catégories générales et en « besoins en évaluation » plus spécifiques.....	203
2.3. Prise en compte des caractéristiques de différents commerces électroniques ..	205
2.4. Définition du format de l'évaluation.....	206
3. Découpage des 13 catégories et identification des besoins en évaluation	207
3.1. Les données.....	207
3.1.1. Une première ébauche des besoins en évaluation des professionnels de la veille-marketing	207

3.1.2. Les regroupements des dimensions utilisateurs par les professionnels de la veille-marketing	207
3.1.2.1. Les regroupements du premier professionnel de la veille-marketing	208
3.1.2.2. Les regroupements du second professionnel de la veille-marketing	209
3.1.3. L’outil d’aide à l’analyse experte de sites de commerce (M.A.S.C.)	210
3.1.4. L’entretien avec le responsable du Webmastering	211
3.1.5. Une trace de l’activité : un questionnaire d’évaluation du Sitemarchand.com.....	211
3.2. Le découpage des 13 catégories et leur contenu	212
3.2.1. Notoriété	213
3.2.2. Facilité d’accès & Capacité du système.....	215
3.2.3. Page d’accueil	216
3.2.4. Offre & Positionnement - Description des produits – Look & co	217
3.2.5. Offre & Positionnement	218
3.2.6. Description des produits	220
3.2.7. Look, présentation, visibilité et navigation.....	221
3.2.8. Netmarketing.....	223
3.2.9. Sécurité & Paiement	224
3.2.10. Interactivité	226
3.2.11. Logistique	227
3.2.12. Les informations associées au service de vente	228
3.2.13. Facilité à trouver un produit.....	229
3.2.14. Facilité à commander un produit	230
4. Conclusion.....	231
CHAPITRE 6	233
L’OUTIL D’AIDE A LA CONCEPTION D’EVALUATIONS POUR LES PROFESSIONNELS DU WEB MARCHAND	233
Repères pour le chapitre.....	235
1. Introduction	237
2. Validation de l’ébauche de la boîte à outils	238
3. Présentation de la boîte à outils validée	239
3.1. Introduction.....	239
3.2. Présentation de la boîte à outils élément par élément	240
3.2.1. Catégorie « Notoriété »	241
3.2.1.1. Notoriété	241
3.2.1.2. Image de marque	241
3.2.1.3. Efficacité et poids relatif des différents supports médiatiques	243
3.2.1.4. Raisons de retour des acheteurs sur le site.....	244
3.2.1.5. Réception des messages publicitaires	245
3.2.2. Catégorie « Facilité d’accès & Capacité système »	246
3.2.2.1. Référencement et liens d’accès au site	246
3.2.2.2. Rapidité de chargement	247
3.2.2.3. Simplicité de l’adresse URL.....	249

3.2.3. Catégorie « Page d'accueil ».....	250
3.2.3.1. Compréhension immédiate de l'objectif du site	250
3.2.3.2. Réassurance	251
3.2.3.3. Rubriques & Services principaux	253
3.2.4. Catégorie « Look, présentation, visibilité et navigation »	254
3.2.4.1. Reconnaissance visuelle	255
3.2.4.2. Attractivité & Esthétique	256
3.2.4.3. Graphismes & Animations	258
3.2.4.4. Aération	258
3.2.4.5. Organisation & Mise en valeur.....	259
3.2.4.6. Boutons d'action.....	261
3.2.4.7. Apparence des courriels.....	262
3.2.4.8. Apparence des « Cross-Sellings » et des Opérations spéciales	263
3.2.4.9. Publicité en ligne	264
3.2.5. Catégorie « Offre & Positionnement »	267
3.2.5.1. Positionnement par rapport à la concurrence.....	267
3.2.5.2. Les motifs d'achats	269
3.2.5.3. Choix de produits – Prix – Promotions & Nouveautés	270
3.2.5.4. Avantages clients.....	271
3.2.5.5. Positionnement interne	272
3.2.6. Catégorie « Description des produits ».....	273
3.2.6.1. Utilité des informations-produits.....	273
3.2.6.2. Caractéristiques descriptives des produits	274
3.2.6.3. Commentaires & Critiques	276
3.2.6.4. Illustrations	277
3.2.6.5. Annonce prix	278
3.2.6.6. Annonce disponibilité & frais d'envoi	279
3.2.6.7. Mise à jour – « Fraîcheur » des produits	280
3.2.6.8. Editorial	281
3.2.7. Catégorie « Netmarketing »	281
3.2.7.1. Esprit communautaire et offres spéciales ou croisées	282
<i>Esprit communautaire</i>	282
<i>Vote électronique & Témoignages</i>	283
<i>Présélections dans les listes de résultats</i>	284
« Cross-Sellings »	285
<i>Opérations & Boutiques spéciales en ligne</i>	286
<i>Promotion « online » des produits</i>	287
<i>Liens vers des sites d'informations, des chats, des forums</i>	289
3.2.7.2. « Newsletters », courriels « en push » et « compte perso »	291
<i>Courriels en « push » & « Newsletters » ou offres personnalisées</i>	291
<i>Compréhension de la page « compte perso »</i>	294

3.2.8. Catégorie « Logistiques & Services ».....	295
3.2.8.1. Utilité des services.....	295
3.2.8.2. Suivi de commande	296
3.2.8.3. Garanties de livraison	297
3.2.8.4. Possibilité de RDV & Relais colis.....	299
3.2.8.5. Emballage cadeau	300
3.2.8.6. Frais d'envoi.....	301
3.2.8.7. Retour de produits	302
3.2.8.8. Service de maintenance	303
3.2.9. Catégorie « Interactivité ».....	304
3.2.9.1. Possibilité de contacter un service clients	305
3.2.9.2. Qualité du service clients.....	307
3.2.10. Catégorie « Sécurité & Paiement ».....	308
3.2.10.1. Facilités de paiement	308
3.2.10.2. Sentiment de sécurité.....	309
3.2.10.3. Confidentialité	310
3.2.10.4. Options de sécurisation.....	311
3.2.10.5. Partenaires financiers et organismes de contrôle.....	313
3.2.11. Catégorie « Facilité à trouver un produit »	314
3.2.11.1. Trouver un produit.....	314
3.2.11.2. Recherche par mots-clefs.....	316
3.2.11.3. Rubriques et sous-rubriques de produits	317
3.2.11.4. Recherche avancée et autres modes de recherche	318
3.2.11.5. Listes de résultats	319
3.2.12. Catégorie « Facilité à commander un produit »	321
3.2.12.1. Commander un produit.....	321
3.2.12.2. Saisie des caractéristiques des produits	322
3.2.12.3. Panier.....	324
3.2.12.4. Bon de commande	327
3.2.12.5. Récapitulatif frais & délais	329
3.2.12.6. Création d'un identifiant & « compte perso »	330
3.2.13. Catégorie « Informations associées au service de vente »	332
3.2.13.1. Compréhension et complétude des informations associées au service de vente	332
4. Conclusion.....	333
CHAPITRE 7	335
LES DEUX VERSANTS DE L'ÉVALUATION ERGONOMIQUE : SYNTHÈSE ET PERSPECTIVES.....	335
1. Synthèse et discussions	337
1.1. Apport de connaissances pour les utilisateurs des techniques d'élucidation.....	337
1.1.1. Concernant les dimensions utilisateurs versus experts	337
1.1.2. Pour l'élaboration de dimensions et d'un questionnaire d'évaluation....	338
1.1.3. Pour produire des résultats utiles à la conception du produit/service.....	339

1.2. Prise en compte des professionnels, producteurs et utilisateurs de résultats d'évaluation.....	339
1.2.1. Différents mouvements en ergonomie par rapport à la prise en compte des concepteurs	339
1.2.1.1. Communiquer les résultats ergonomiques.....	340
1.2.1.2. Faire participer les concepteurs à l'évaluation	340
1.2.1.3. Donner aux concepteurs leurs propres moyens d'évaluation	341
1.2.1.4. Prendre en compte l'activité des concepteurs.....	342
1.2.1.5. La position adoptée dans cette étude	342
1.2.2. Les particularités de la boîte à outils (en prenant en compte l'activité des sujets de type 1, 2 et 3).....	342
1.2.3. Développement d'instruments futurs probables par les sujets de type 1 et 2 du Web marchand	347
1.2.3.1. Transformation matérielle de l'artefact en vue de former un instrument assurant une médiation pragmatique	348
<i>Instrumentalisation par réduction du champ des possibles</i>	348
<i>Instrumentalisation par ajout et modification de composants de l'artefact</i> .	349
1.2.3.2. Attribution d'autres fonctions à la boîte à outils	350
<i>Un instrument de type check-list pour les sujets de type 1 ou 2 (fonction réflexive)</i>	350
<i>Un instrument de développement de connaissances sur les possibilités d'évaluation pour les sujets de type 1 et 2 (fonction épistémique)</i>	351
1.2.3.3. Instrumentation.....	351
1.2.3.4. Le rôle de l'ergonome, encore nécessaire ?.....	352
2. Pour conclure, d'autres recherches nécessaires... ..	353
2.1. Etudier l'activité des utilisateurs de produits/systèmes et celle des concepteurs comme deux axes de recherche complémentaires	353
2.2. Ergonomie et marketing.....	354
2.3. Etudier l'activité d'élaboration d'instruments à partir de l'artefact « boîte à outils »	355
 BIBLIOGRAPHIE	 357
 ANNEXES	 387
Annexes du chapitre 3 :.....	389
Guide de l'entretien individuel	389
<i>Accueil</i> :	389
Phase de verbalisation sans utilisation du produit/service.....	389
1. <i>Technique d'élicitation associative</i> :	389
2. <i>Entretien d'explicitation</i> :	389
3. <i>Phase de synthèse écrite (fiches-critère)</i> :	389
Phase de verbalisation avec utilisation du produit/service	390
4. <i>Méthode des verbalisations libres</i> :	390
5. <i>Phase de synthèse écrite (seconde liste de fiches-critère)</i> :	391
Guide du <i>focus group</i>	392
<i>Accueil</i> :	392

Phase de verbalisation sans utilisation du produit/service.....	392
1. (et 3.) <i>Technique d'élicitation associative et création et nomination de critères collectifs</i> :	392
2. <i>Discussions à partir d'expériences vécues</i> :	393
3. <i>L'étape 3</i> :	393
Phase de verbalisation avec utilisation du produit/service	393
4. <i>La méthode des verbalisations libres</i> :	393
5. <i>Discussion de synthèse sur les critères de jugements</i> :	394
Liste exhaustive de dimensions utilisateurs élaborées à partir des données des entretiens individuels et des <i>focus groups</i>	395
Ressources humaines nécessaires et coûts de réalisation de différentes étapes de l'entretien individuel et du <i>focus group</i>	396
 Annexes du chapitre 4 :	 399
Entretien avec un ingénieur concepteur de FTA commentant le rapport d'évaluation de la satisfaction des utilisateurs d'une application logicielle pour l'accès à Internet :	399
Résumé des besoins en connaissances des concepteurs révélés lors de cet entretien :	399
Description sommaire de l'étude d'évaluation commentée.....	400
Commentaires du concepteur quant aux résultats d'évaluation.....	401
Commentaires par rapport aux résultats concernant la présentation multimédia .	401
Commentaires par rapport aux résultats concernant le processus d'installation et l'aide.....	402
Commentaires par rapport aux résultats concernant l'interface avec boutons d'accès aux principales fonctionnalités	404
Commentaires par rapport à la conclusion de l'étude d'évaluation.....	404
Extrait de la liste exhaustive de dimensions utilisateurs de sites Internet marchands avec leurs sous-dimensions, avant la transformation de surface	405
Liste de dimensions utilisateurs de sites Internet marchands transformée en surface, avant l'entretien de validation avec le responsable du Webmastering	406
<i>Dimensions d'évaluation liées plus particulièrement aux sites Internet marchands</i> :	406
<i>Dimensions d'évaluation génériques, relatives à tout site Internet</i> :	410
Liste de dimensions utilisateurs de sites Internet marchands transformée en surface, après l'entretien de validation avec le responsable du Webmastering	411
<i>Dimensions d'évaluation liées plus particulièrement aux sites Internet marchands</i> :	411
<i>Dimensions d'évaluation génériques, relatives à tout site Internet</i> :	414
 Annexes du chapitre 5.....	 415
Les regroupements des dimensions utilisateurs des sites Internet marchands par le premier professionnel de la veille-marketing (spécialiste des sites marchands)	415
Les regroupements des dimensions utilisateurs des sites Internet marchands par le second professionnel de la veille-marketing (non spécialiste des sites marchands)	418

Annexes du chapitre 6.....	420
Récapitulatif des « besoins en évaluation » avant et après validation de l'ébauche de l'outil d'aide à la conception d'évaluations	420
« Notoriété perçue », avant et après validation.....	420
« Facilité d'accès & Capacité système », avant et après validation	420
« Page d'accueil », avant et après validation	421
« Look, présentation, visibilité et navigation », avant et après validation...	421
« Offre & Positionnement », avant et après validation	421
« Description des produits », avant et après validation	421
« Netmarketing », avant et après validation.....	422
« Logistique & Services », avant et après validation	422
« Interactivité », avant et après validation	422
« Sécurité & Paiement », avant et après validation	423
« Facilité à trouver un produit », avant et après validation.....	423
« Facilité à commander un produit », avant et après validation.....	423
« Informations associées au service de vente », avant et après validation ..	423

INTRODUCTION GENERALE

Cette thèse se situe dans le champ de l'évaluation ergonomique de produits et de services. Une bonne évaluation doit répondre à deux types de critères qui ne sont pas nécessairement convergents. D'une part, il s'agit de produire des résultats qui répondent au critère de vérité de la connaissance. D'autre part, il s'agit de produire des résultats qui répondent au critère pragmatique de l'efficacité et de l'utilité : un « bon » résultat d'évaluation du point de vue d'un ergonomiste est-il aussi nécessairement « bon » du point de vue des professionnels qui vont l'utiliser pour transformer le produit/système ? Qu'est-ce qu'un « bon » résultat ergonomique : celui qui relate une connaissance très précise de l'utilisation et de la perception des produits/systèmes ou celui qui permet une prise de décisions pour la transformation et l'amélioration du produit/système ? Dans la thèse, nous allons explorer les deux dimensions. Nous réalisons d'abord une étude pour les professionnels de l'évaluation de France Télécom en adoptant leur point de vue orienté vers la production de connaissances sur l'utilisateur. Dans un second temps, nous réalisons un travail à partir d'entretiens menés avec les professionnels d'un service de veille-marketing sur les nouvelles technologies et d'une entreprise marchande par Internet, afin d'étudier les exigences de leur activité de transformation des produits/services.

Dans la première étude, nous cherchons à répondre à la demande des professionnels du service d'évaluation de France Télécom¹, pris dans un contexte de compétition industrielle croissante ; cela, en raison de la conjoncture générale (les temps de mise au point d'un produit sont de plus en plus réduits ; Desurvire, 1994; Gaillard, 1997) et en raison de l'ouverture de l'entreprise au marché libre. Ces professionnels de l'évaluation partagent la préoccupation de nombreux ergonomistes qui est de développer les méthodes d'évaluation les plus efficaces et les plus économiques possibles, tout en préservant la part essentielle de la validité des données ergonomiques. Les professionnels de France Télécom élaborent des évaluations par questionnaire et ont besoin d'éléments de comparaison sur deux techniques utilisées pour recueillir les critères de jugement des utilisateurs (dimensions utilisateurs) et un contenu verbal aidant la formulation du questionnaire : l'entretien individuel et le *focus group*². Nous menons une étude comparative de ces techniques dans le domaine des sites Internet marchands.

¹ Où nous avons été accueilli dans le cadre d'un contrat à durée déterminée de thèse durant trois ans.

² Le *focus group* est une discussion de groupe, avec 5 à 10-12 participants, modérée par un animateur qui doit centrer (*focus*) les discussions sur les informations recherchées.

Plusieurs questions orientent ce travail :

- En quoi se distinguent des dimensions d'évaluation élaborées à partir de données recueillies auprès d'utilisateurs (dimensions utilisateurs) et des dimensions d'évaluation élaborées par des experts à partir de leurs connaissances des utilisateurs/clients (dimensions experts) ?
- Comment se distinguent l'entretien individuel et le *focus group* en termes de nature et d'exploitabilité du contenu verbal produit et en termes de productivité (production verbale par rapport aux coûts humains et financiers de réalisation) ?
- A partir de quel nombre de sujets, la productivité des techniques commence-t-elle à stagner ou à décliner ?
- Quel est l'impact de l'utilisation (versus non utilisation) du produit/service sur la nature et l'exploitabilité de la production verbale ?

Nous mettons ensuite en évidence deux points de vue différents sur la question de la pertinence d'une évaluation, portés par les professionnels de l'évaluation, d'une part, et par ceux du marketing, d'autre part.

- Pour les professionnels de l'évaluation, seules les dimensions utilisateurs sont pertinentes à évaluer. Ils sont tournés vers une logique de production de connaissances (sur les utilisateurs). Ce point de vue est lié, en partie, à leur activité dédiée exclusivement à l'évaluation ergonomique. Le service d'évaluation est « en bout de chaîne » et traite l'évaluation des produits/services conçus par d'autres services et branches opérationnelles de France Télécom.
- Pour les professionnels du marketing, seuls les résultats dont ils peuvent effectivement se servir pour transformer leurs produits/services doivent être produits. Ils sont engagés dans un mouvement pragmatique.

D'un côté, les professionnels de l'évaluation développent un processus d'instrumentalisation des dimensions utilisateurs pour construire le questionnaire d'évaluation et se construisent ainsi un instrument d'élaboration d'évaluations à partir de cet artefact. D'un autre côté, les professionnels du marketing ne parviennent pas à instrumentaliser cet artefact, alors qu'ils réalisent eux-mêmes des évaluations par questionnaire de leurs produits/services. Dans cette seconde situation, les professionnels ne parviennent pas à se construire un instrument à partir de l'artefact.

Les professionnels de l'évaluation tiennent compte de la relation aux artefacts, seulement du point de vue de l'activité des utilisateurs. Or, s'il s'agit de traiter la relation aux artefacts comme problème central, cette dernière doit être explorée selon deux axes correspondant à deux formes d'activités humaines qui ne peuvent être pensées indépendamment : celles des utilisateurs des produits/services, considérés avec leur culture et leur histoire dans leur contexte singulier, et celles des divers acteurs de la conception (Bannon & Bodker, 1991). En effet, l'ergonome doit s'intéresser de plus près à l'activité des concepteurs³ pour leur fournir des aides réelles. L'objectif est que les résultats d'évaluation soient des outils pour les professionnels transformant les produits/services.

Nous nous intéressons à l'activité des professionnels du marketing, mais aussi à celle des professionnels de la conception du Web marchand pour produire, non pas directement des résultats, mais un outil d'aide à la production de résultats pour eux. Nous construisons un outil qui vise à être un instrument pour outiller leur activité d'identification de leurs besoins en connaissances, c'est-à-dire ce qu'ils ont besoin de savoir à l'issue d'une évaluation, et leur activité d'élaboration de cette évaluation. Pour cela, notre travail consiste à opérationnaliser les précédentes dimensions utilisateurs en tenant compte de leurs représentations, des exigences de leur activité et donc de leurs besoins en connaissances (relatifs aux aspects ergonomiques, commerciaux et logistiques d'un commerce électronique). L'outil ou plutôt la boîte à outils construite formalise ces besoins et propose des techniques d'évaluation adaptées.

Sur un plan pratique, la demande initiale du service d'évaluation et notre situation de salarié au sein de ce dernier ont occasionné des difficultés pour étudier l'activité des professionnels de la transformation. La seconde partie de cette thèse correspond ainsi au fruit d'un travail réalisé, en partie, en dehors de l'entreprise et il n'a pas été possible d'étudier l'activité d'utilisation de « réels » résultats d'évaluation.

Ce texte est organisé de la façon suivante :

- Le chapitre 1 établit l'état de l'art des méthodes et techniques d'évaluation. Nous y relatons les préoccupations en termes d'évaluation de diverses époques et disciplines, depuis l'apparition du *design* d'un produit et de sa communication, au milieu du siècle dernier, jusqu'à l'émergence récente d'une nouvelle forme d'évaluation, subjective et sensorielle. En

³ Dans le terme générique de « concepteur », nous comprenons aussi bien les professionnels qui réalisent techniquement le produit/service, que les professionnels qui s'occupent de l'aspect plus commercial du projet. Les concepteurs et « marketers » sont ainsi rassemblés sous le terme de « professionnels de la transformation ».

matière d'évaluation ergonomique, nous distinguons les techniques faisant intervenir l'utilisateur des techniques d'évaluation expertes. Parmi les premières, nous différencions celles basées sur le recueil de données expérimentales, celles basées sur le recueil de données verbales et enfin celles basées sur le recueil de données subjectives (hédoniques).

- Le chapitre 2 expose les questions de recherche relatives (a) à l'élaboration de dimensions utilisateurs versus experts, (b) à la comparaison de l'entretien individuel et du *focus group*, (c) au nombre de sujets nécessaires pour recueillir un ensemble suffisant de données et (d) à la comparaison de la situation de verbalisation avec versus sans utilisation. Nous abordons ensuite le renversement de point de vue porté sur l'évaluation. La nécessité d'étudier conjointement l'activité des utilisateurs de produits/services et celle des concepteurs étant au cœur de l'approche instrumentale (Rabardel, 1995), nous l'utilisons et la présentons.

- Le chapitre 3 présente l'étude relative aux premières questions de recherche. Les résultats nous conduisent à discuter de la pertinence de l'entretien individuel et du *focus group* en fonction du type d'évaluation par questionnaire à construire et des conditions de réalisation données (ex. niveau de connaissances des professionnels). Nous développons aussi des arguments sur l'utilité de certains résultats qualitatifs pour la conception et proposons des pistes d'usage des deux techniques, avec et sans utilisation de produits/services, pour la construction d'un questionnaire comme pour la conception. Une indication est enfin donnée concernant le nombre de sujets au-delà duquel la productivité des techniques décline.

- Le chapitre 4 met en perspective les deux points de vue sur l'évaluation. Nous retraçons d'abord le processus mis au point par les professionnels de l'évaluation, organisé autour d'une liste exhaustive de dimensions utilisateurs. Nous menons ensuite des entretiens avec cinq professionnels, de la veille-marketing et de la conception de sites Internet (marchands), pour étudier leur compréhension de ces dimensions utilisateurs. Nous transformons alors ces dernières, en surface, puis montrons en quoi, même transformées, elles ne peuvent être instrumentalisées par les professionnels de la veille-marketing. Cette analyse débouche sur la nécessité de concevoir un artefact instrumentalisable pour les aider dans leur propre activité d'évaluation.

- Dans le chapitre 5, nous spécifions le contenu (les besoins en connaissances des professionnels) et les caractéristiques structurelles et fonctionnelles du futur outil d'aide à la conception d'évaluations destiné à tout professionnel du Web marchand et, dans le chapitre 6, nous décrivons la boîte à outils construite, élément par élément, après avoir fait valider sa première ébauche par quatre professionnels du Web marchand (marketing et conception).

- Enfin, dans le chapitre 7, nous explicitons l'intérêt de l'artefact conçu, en rapport avec les activités des évaluateurs, des utilisateurs des produits/services et des utilisateurs des résultats d'évaluation. Nous décrivons le développement d'instruments futurs probables à partir de la boîte à outils, d'une part, et l'intérêt d'un mouvement d'échanges et de recherche de compromis entre les deux groupes de professionnels (évaluation versus transformation), d'autre part.

CHAPITRE 1

HISTORIQUE ET ETAT DE L'ART DES METHODES D'EVALUATION DE PRODUITS, SYSTEMES ET SERVICES

1. INTRODUCTION

Il existe une littérature très abondante sur les méthodes et techniques d'évaluation de produits, de services, de systèmes informatiques ou outils de travail. On trouve également de nombreuses classifications pour organiser cette profusion. On peut citer la distinction entre techniques prédictives et techniques expérimentales (Coutaz, 1994) ou entre approches analytiques et approches empiriques (Sénach, 1990). Mentionnons également la classification en fonction des composants intervenant dans l'évaluation (un expert et un système ; un modèle de l'utilisateur et un modèle de l'interface; etc.; Howard & Murray, 1987) ou celle en fonction de la présence ou non de l'utilisateur et de la présence ou non du système (Whitefield *et al.*, 1991). Les facteurs relatifs au choix opérationnel d'une méthode d'évaluation (ex. les connaissances requises pour sa mise en œuvre, les ressources matérielles, les facteurs situationnels, les moyens humains et les résultats fournis) peuvent aussi servir à ce découpage (Balbo, 1994) ou encore simplement le nombre de tâches et d'utilisateurs (1 ou n) permettant une vue plus globale (Farenc, 1997).

Toutefois, la classification la plus courante sans doute est fondée sur la distinction entre *formative evaluation* (évaluation au cours du développement du produit) et *summative evaluation* (évaluation d'une solution de conception aboutie) introduite dès la fin des années 1960 (Carroll *et al.*, 1992). Les méthodes peuvent aussi être différenciées en fonction d'un plus grand nombre d'étapes dans le processus de conception : celles utilisées pour renseigner le problème de conception, lors du développement du prototype, lors de l'évaluation de la solution de conception et éventuellement lors de l'évaluation de « feed-back » après son lancement sur le marché (Baccino *et al.*, 2005; Bastien & Scapin, 2004; Stanton & Baber, 1996b). Considérant l'évaluation, tenir compte de la phase amont à la conception peut sembler surprenant. Cependant, il y a une progression de l'évaluation ergonomique dans le sens de renseigner davantage les problèmes de la conception au fur et à mesure de ses itérations et d'abandonner l'évaluation (expérimentale et statistique) d'une solution de conception déjà formée (Karat, 1997). Cette évolution peut permettre de prendre davantage en considération l'usage réel de l'artefact au cours de son développement et de contribuer à prendre pour objet de recherche principal l'utilisation de l'artefact et non l'artefact lui-même (Bannon & Bodker, 1991). Nous nous inscrivons dans la même perspective et abordons donc diverses méthodes et techniques intervenant dans toute la « boucle » de conception, même si nous nous arrêtons davantage sur les méthodes d'évaluation d'une solution en cours d'élaboration ou déjà aboutie.

De plus, nous choisissons une approche historique et par discipline pour l'exposé de cet état de l'art. Cette approche n'a pas été entreprise jusqu'à présent ou du moins pas de façon aussi complète. La discipline d'origine de chaque méthode ou technique et son point de départ dans l'histoire (préoccupations sociales et scientifiques du moment) présentent pourtant une importance non négligeable pour leur compréhension et leur utilisation adéquate. Il nous semble qu'un recueil de l'ensemble de ces repères est manquant.

Par ailleurs, nous choisissons d'embrasser un assez grand nombre de disciplines, des sciences humaines aux diverses approches de la conception, en passant par le marketing. Le court exposé de Macaulay (1996) en est un bon exemple. Après avoir brossé les débuts du marketing avec l'apparition du *design* d'un produit et de sa communication, nous abordons la conception participative. Courant de conception plus démocratique, elle adapte des techniques de l'ethnographie pour observer l'activité des opérateurs sur le terrain et utilise aussi des techniques de créativité et le *focus group* auquel nous nous intéressons particulièrement dans cette thèse. Nous examinons ensuite le recueil de données expérimentales, qu'il s'agisse d'une expérimentation contrôlée ou d'un test utilisateur moins formel, puis nous abandonnons les données comportementales pour étudier la lente habilitation des données verbales et de la notion de subjectivité dans la psychologie en général et dans l'évaluation des produits en particulier. Parallèlement à ce mouvement, dans les années 1990, un accroissement de la compétition industrielle engendre le recours à des techniques d'évaluation plus rapides, sans l'utilisateur. Nous les décrivons et les critiquons. Cette époque voit aussi débiter l'évaluation subjective et sensorielle : les produits doivent être évalués sur des dimensions hédoniques et émotionnelles auxquelles les utilisateurs et clients sont désormais sensibles. Nous détaillons les techniques de cette nouvelle forme d'évaluation appliquée à divers domaines industriels, ainsi que les enquêtes par questionnaire qui permettent, entre autres, la prise en compte de ces dimensions dans l'évaluation ergonomique.

2. HISTORIQUE ET ETAT DE L'ART DES METHODES ET TECHNIQUES D'EVALUATION

Certains facteurs d'après-guerre ont favorisé le développement de l'ergonomie puis l'émergence d'une conception centrée sur l'utilisateur. Cette approche anthropocentrique trouve ses origines scientifiques dans l'école de la psychologie russe et est accompagnée d'une meilleure intégration de l'ergonomie au niveau industriel (dans tous les pays dits

développés). Nous parcourons rapidement cette évolution car elle est accompagnée du développement de méthodes et techniques d'évaluation décrites dans ce chapitre.

La période de reconstruction d'après-guerre se traduit par une automatisation de l'industrie entre 1950 et 1975. Il s'agit de produire en masse pour une demande importante. Ces transformations se réalisent « *selon des schémas de penser technologiques et tayloriens* » (Jardillier, 1992, p. 13). Le Taylorisme venu des Etats-Unis, permet d'organiser cette production de masse. Basé sur le découpage du travail en menues opérations et sur la division des fonctions d'organisation et d'exécution du travail, il engendre une déshumanisation. Cela se traduit par des dysfonctionnements sociaux, une baisse de la qualité de la production et un accroissement des accidents du travail. Le Taylorisme et plus généralement l'évolution sociale et technique contribuent également à séparer de plus en plus les fonctions de conception, de fabrication et d'utilisation des produits ou outils (Gaillard, 1997). Enfin, cette ère de progrès scientifique et technique est aussi ponctuée de catastrophes (ex. catastrophe de Tchernobyl) qui mettent au jour des erreurs inacceptables. Ces évolutions ont favorisé le développement de l'ergonomie. D'une part, les ergonomes vont s'intéresser davantage à l'activité d'organisation des tâches d'exécution, d'autre part, il s'agit de prendre en compte le concept de sûreté de fonctionnement dans la conception (Gaillard, 1997). Les ergonomes vont considérer de plus près l'activité de conception des ingénieurs, des architectes, des informaticiens, etc. à qui il s'agit de transmettre les résultats relatifs à l'activité de l'utilisateur pour une nécessaire conception centrée sur l'utilisateur. Ainsi, une approche anthropocentrique, permettant de replacer l'homme et son activité au cœur de la démarche et finalité de conception, s'est développée à partir des années 1980 (Rabardel, 1995).

Sur le plan scientifique, cela trouve son origine dans l'approche anthropocentrique adoptée par la psychologie de l'ingénierie⁴ soviétique qui se développe sous la direction de Lomov (1975, 1977 - cité par Nosulenko & Samoylenko, 1999a). En URSS, dès les années 1960-1970, la psychologie acquière une place universitaire à part entière aux côtés des sciences dites « dures » et un statut intégrateur pour les sciences naturelles, techniques et humaines, dans le cadre des sciences de l'homme où des recherches pluridisciplinaires sont menées (Nosulenko & Samoylenko, 1999a). Cela permet de développer une approche anthropocentrique qui s'oppose à l'approche technocentrique représentant l'homme en tant qu'un simple élément du système « homme-machine » (Rabardel, 1995).

⁴ Elle se définit ainsi comme une science à l'intersection des disciplines biologiques, techniques et anthropocentriques et constitue un équivalent des disciplines des Human Factors développés aux USA.

Sur le plan social, c'est-à-dire au niveau industriel, Gould et Lewis (1985) sont parmi les premiers à proposer une méthodologie de « conception centrée utilisateurs » (*designing for usability*). Les auteurs prêchent pour une prise en compte anticipée de l'utilisateur dans le processus de conception et dénoncent l'idée préconçue selon laquelle les itérations de conception sont seulement un affinement coûteux. Ils argumentent que le produit sera, de toute façon, évalué lors de son achat et que les modifications seront plus coûteuses si réalisées à partir de ce stade. D'ailleurs, l'approche anthropocentrique⁵ a aussi pu se mettre en place concrètement grâce à la prise de conscience de la diminution des coûts de conception (Balbo, 1994; Coutaz, 1994) et de l'augmentation du profit réalisé⁶ (Baccino *et al.*, 2005; Griffin & Hauser, 1993) avec une prise en compte anticipée de l'utilisateur/client.

Cette approche est encadrée par le concept d'« utilisabilité » qui émerge également dans les années 1980. Le concept s'affinera à travers diverses contributions avant d'aboutir à sa définition finale dans les normes internationales ISO 9241-11 (ISO, 1998 ; efficacité, efficacité, satisfaction). Sa formalisation a encouragé la prise en considération de critères d'utilisabilité (Jordan *et al.*, 1996b), même si certains affirment que ce concept renvoie encore trop souvent à des significations variées (Stanton & Baber, 1996b).

Par ailleurs, l'ergonomie est de plus en plus intégrée dans le processus de conception et les commanditaires des études de plus en plus sensibilisés à l'ergonomie, dont ils ont constaté l'apport en termes de valeur ajoutée dans leurs produits (Jordan *et al.*, 1996a; Jordan *et al.*, 1996b; Nielsen, 1993). Des réserves sont toutefois émises quant à cette évolution. Par exemple, selon Mayhew (1999), une méthodologie de « conception centrée utilisateurs » resterait à être conçue et institutionnalisée de façon plus vaste pour, non pas insérer de l'ergonomie dans le processus de conception, mais pour concevoir ce processus autour des questions de l'utilisabilité. D'autre part, les concepteurs connaissent toujours peu les principes basiques de la « conception centrée utilisateurs ». Dans leur étude de 1985, Gould et Lewis avaient interrogé 447 concepteurs informaticiens sur les étapes à suivre pour développer et évaluer un nouvel ordinateur pour des utilisateurs finaux, et montraient qu'environ un tiers seulement donnait une réponse en rapport avec l'un de leur trois fondamentaux⁷. La situation serait toujours la même quinze ans plus tard (Rubin, 1994). Malgré ces remarques plus négatives, nous constatons que le nombre d'entreprises équipées

⁵ Plus connue, dans le milieu industriel, sous la notion de « qualité totale » où les entreprises développent des techniques pour connaître l'ensemble des besoins utilisateurs/clients et concevoir leur produits/services à partir de ces derniers.

⁶ Corrélatif à l'augmentation de la satisfaction du consommateur comme à celle des employés de l'entreprise.

⁷ « *Early focus on users and tasks* », « *Empirical measurement* », « *Iterative design* ».

de laboratoire de test ergonomique et la littérature intarissable sur les modèles *user centered design* témoignent d'une évolution certaine.

Ce développement de l'intégration de l'ergonomie dans le processus de conception est accompagné d'un développement foisonnant de méthodes et techniques d'évaluation, parfois issues de l'adaptation de méthodes développées par d'autres disciplines. Avant de décrire cette profusion de méthodes ergonomiques qui datent des années 1980-1990, nous décrivons l'apparition du *design* d'un produit et de sa communication qui s'accompagnent d'une certaine forme d'évaluation.

2.1. L'EVALUATION QUALITATIVE EN MARKETING ET EN SCIENCES SOCIALES APPLIQUEES

2.1.1. Evaluer le design d'un produit et sa communication

Le boom industriel d'après-guerre et la compétition qui l'accompagne obligent les professionnels du marketing à chercher comment les produits de leur entreprise peuvent être rendus plus attractifs, et cela particulièrement aux Etats-Unis (Krueger & Casey, 2000). Ils cherchent à connaître les valeurs et opinions des consommateurs pour adapter leurs produits, emballages ou encore publicités aux dernières tendances sociales observées. En effet, ils s'aperçoivent que les gens consomment parfois un produit car il véhicule une valeur sociale positive. Ils évaluent donc si leurs produits et la façon de les « marketer » correspondent aux attentes et aux représentations des consommateurs.

Dans cette approche qualitative de marketing, ils utilisent, d'une part, l'entretien approfondi et, d'autre part, le *focus group* qui est basé sur des discussions de groupe (Ives, 2003). Le *focus group* y est largement privilégié malgré les réels avantages de l'entretien approfondi pour certains objectifs d'étude (Gordon & Langmaid, 1988). Toutes les grandes entreprises des Etats-Unis sont dotées de leur salle d'entretien à glace sans tain permettant aux clients de l'étude d'observer le déroulement d'un *focus group* (Krueger & Casey, 2000). Ce dernier est apprécié des professionnels du marketing car, selon eux, il permet d'obtenir des résultats crédibles, rapides et peu coûteux.

Pendant de nombreuses décennies, ces techniques (mais d'autres aussi comme le *brainstorming*, les entretiens par paire ou par trio, etc.) vont permettre aux entreprises d'adapter le *design* de leurs produits aux différents segments de consommateurs, d'inventer

de nouveaux produits, de réviser le positionnement de leur marque (mise en exergue des valeurs qu'elle véhicule par rapport à celles des concurrents) ou de modifier leurs campagnes publicitaires (Gordon & Langmaid, 1988).

2.1.2. L'émergence du focus group avec l'intérêt porté à la recherche qualitative

Pourtant, durant les décennies 1960-1970, la recherche qualitative connaît des difficultés. L'époque fait davantage confiance aux chiffres et est plus tournée vers des études quantitatives statistiques en marketing comme en sciences sociales appliquées⁸ pour, par exemple, mettre en évidence les préférences d'une large population (Ives, 2003). Ces études reposent alors sur des dimensions définies a priori et le problème est que ces dernières sont issues des hypothèses élaborées par le chercheur ou le professionnel du domaine. Elles reflètent donc sa vision du monde et pas forcément le point de vue des sujets (Krueger & Casey, 2000; Merton *et al.*, 1956). Néanmoins, les préoccupations du moment portent davantage sur les procédures expérimentales, l'élaboration des échantillons de sujets et des groupes de contrôle, en particulier en sciences sociales. C'est seulement à partir des années 1970, et surtout durant les années 1980, que les débats concernant la supériorité de la recherche quantitative sur la recherche qualitative, et inversement, s'estompent (Gordon & Langmaid, 1988; Ives, 2003; Morgan, 1993). Ils laissent enfin place à des études qui ne soient plus défensives mais qui cherchent à mieux établir les avantages et inconvénients de chaque type de recherche et de ses méthodes (Morgan, 1993).

Les chercheurs en sciences sociales s'intéressent à nouveau au *focus group* qu'ils ont rejeté une quarantaine d'années auparavant. Pourtant, dès les années 1930, ils ont des doutes sur leurs techniques d'entretien utilisant des questionnaires à choix multiples qui contraignent les répondants dans des réponses « toutes faites », et des techniques moins directives sont alors explorées (Krueger & Casey, 2000). Merton (Merton *et al.*, 1956; Merton & Kendall, 1946) met au point une méthode d'appréciation des effets sociaux et psychologiques des moyens de communication de masse sur le grand public. Le principe est que des participants, recrutés à cet effet, visualisent un film ou écoutent l'enregistrement d'un programme de radio et participent ensuite à des entretiens non directifs, ciblés (*focus*) sur les événements qui ont

⁸ Où il s'agit, par exemple, d'évaluations sur le terrain suite au déploiement d'un programme de réformes sociales ou encore de connaître la compréhension du grand public quant à la maladie mentale, aux causes d'une attaque cardiaque, etc.

retenu leur attention pour comprendre leurs sentiments et opinions⁹. Il s'agit de *Focused Interviews* qui sont l'ancêtre du *focus group*.

A cette époque, les chercheurs en sciences sociales ne sont pas les seuls à s'intéresser à nouveau au *focus group*. Dès les années 1980, les ergonomes se sont aussi emparés de cette technique pour aider la conception de dispositifs de travail (Langford & McDonagh, 2003a), mais son utilisation dans ce domaine s'intensifie surtout à partir des années 1990.

Nous décrivons certains de ces travaux les plus récents au paragraphe suivant, lequel retrace l'histoire et les méthodes de la conception participative, un courant de conception qui se veut plus démocratique. Il est accompagné de l'émergence des premières méthodes d'évaluation ergonomiques.

2.2. LA CONCEPTION (ET L'EVALUATION) PARTICIPATIVE

Dans le courant de la conception participative, né dans les années 1970 en Scandinavie, c'est non plus le consommateur mais l'utilisateur et même l'homme comme acteur dans sa situation naturelle d'activité qui est considéré. Différents degrés d'implication de l'utilisateur dans le projet de conception peuvent être envisagés (simple recueil d'informations concernant son expérience de son activité ou recueil de son avis concernant les modifications prévues), le plus haut degré correspondant à son intégration dans les prises de décisions sur l'organisation et la conception des outils de travail (Darses & Reuzeau, 2004).

Ce mouvement prend source dans le « Bauhaus » et le « Nordic Design » qui remontent respectivement aux années 1920 et 1930 (Ehn & Badham, 2002). Il s'agit de mouvements de conception dont la volonté est d'être plus démocratique. Ils cherchent à élaborer des architectures et objets de la vie quotidienne qui soient transparents en termes de fonctionnalités, qui favorisent les échanges sociaux et finalement, dont les propriétés rendent les gens heureux et satisfaits. C'est dans ce contexte qu'est né le *participatory design*.

Dès les années 1970, un projet pionnier associe le Syndicat des Travailleurs de la Métallurgie de Norvège et les chercheurs du « Norwegian Computer Center » et a pour objectif de démocratiser la conception et l'utilisation de la technologie de l'information au travail. La stratégie consiste à mettre en place une production participative de connaissances

⁹ Des informations peuvent également être trouvées concernant les premiers travaux de Merton et leur application dans des recherches pour l'armée américaine dans Stewart et Shamdasani (1990) et Krueger et Casey (2000).

qui puissent être utilisées activement durant la journée de travail dans les usines et au niveau des syndicats. Ce projet débouche sur la mise en place d'accords qui établissent des principes de discussions entre dirigeants et délégués du personnel. En outre, l'objectif est que ces derniers aient accès à des informations sur les systèmes informatiques et sur l'influence de ces systèmes sur les travailleurs (Ehn & Badham, 2002). Ces accords ont ainsi servi à réguler la conception et l'introduction des systèmes informatiques.

Le projet a aussi influencé les stratégies syndicales et même engendré d'importants changements au niveau des lois du travail de la Scandinavie. Un autre projet très connu a également contribué à cela, dans les années 1980 : UTOPIA (Bodker *et al.*, 1993). Basé sur le principe selon lequel la technologie doit être conçue à partir des pratiques du travail, il cherche à favoriser le développement des compétences des opérateurs. Ce projet contribue au développement de nouvelles méthodes basées sur la création de maquettes et de prototypes associées à divers jeux de conception. Nous en détaillons certaines dans ce qui suit.

Cependant, si la conception participative a connu un succès international, surtout au niveau académique, son application au niveau industriel et sa considération dans les sphères politiques ne reflètent pas les objectifs élaborés initialement (Ehn & Badham, 2002). Il semble que le pouvoir réel soit demeuré, malgré tout, très managérial dans les diverses actions participatives (Bodker *et al.*, 1993; Gronbaek *et al.*, 1993). Selon Ehn et Badham (2002), ce qui reste de ce mouvement concerne surtout des méthodes et techniques avec une orientation ethnographique, en particulier dans le domaine des systèmes informatiques.

Nous nous y intéressons en distinguant celles qui concernent :

- l'analyse de la situation à re-concevoir (analyse des caractéristiques des utilisateurs et de leurs besoins relatifs aux tâches de leur activité de travail),
- la conception de solutions comprenant généralement une activité d'évaluation (éventuellement informelle) de ces solutions,
- l'évaluation des solutions conçues.

Même si nous nous intéressons davantage à l'évaluation d'une première forme de solution de conception, à notre sens, ces méthodes forment un continuum et s'appréhendent donc ensemble. Dans la première catégorie, la solution existante est évaluée pour s'orienter vers la conception d'une nouvelle. Dans la suivante, un embryon de solution est constitué et généralement évalué selon une technique prévue ou non. Dans la dernière, enfin, il s'agit de l'évaluation à proprement parler.

2.2.1. Méthodes et techniques d'analyse de la situation

Une technique remarquable contribuant à l'analyse d'une situation donnée est l'enquête contextuelle (*contextual inquiry* ; Holtzblatt & Jones, 1993) faisant partie d'une méthodologie plus générale (*contextual design* ; Holtzblatt & Beyer, 1993). Inspirée des techniques ethnographiques, l'enquête contextuelle consiste à observer les utilisateurs dans leur activité de travail en situation naturelle et à les interrompre si nécessaire pour leur poser des questions. Il est recommandé que des concepteurs mènent ces observations et entretiens car leur implication dans ce recueil de données contribue à générer les meilleurs résultats (Holtzblatt & Beyer, 1993). Cette technique est particulièrement indiquée s'il s'agit de tâches cognitives complexes (ex. dans le domaine médical, financier, etc.; Baccino *et al.*, 2005).

Des activités de « simulation langagière » en groupe peuvent aussi être développées (Daniellou, 2004). On peut citer les *future workshops* (Kensing & Madsen, 1993) dans lesquels un groupe de futurs utilisateurs critique une situation de travail existante, avant d'en imaginer une meilleure sur la base d'activités de dessin et d'exploration de métaphores (« et si la bibliothèque était un magasin... »). Une discussion a ensuite lieu sur les changements nécessaires pour atteindre la situation imaginée. Cette technique peut aussi employer ponctuellement une autre technique participative, celle du *brainstorming*. Cette dernière peut être utilisée pour éliciter les problèmes d'utilisation d'un produit/système et imaginer les nouvelles fonctionnalités souhaitables (Baccino *et al.*, 2005). Enfin, dans les *storytelling workshops* (Greenbaum & Madsen, 1993), deux utilisateurs sont invités à amener deux histoires, l'une positive, l'autre négative, concernant leur activité de travail informatisé et à en commenter les similitudes et différences à travers des discussions de groupe. Cette technique permet de clarifier les problèmes et difficultés des utilisateurs.

2.2.2. Méthodes et techniques de conception de solutions

Une fois les problèmes identifiés, une solution d'amélioration précise peut être recherchée. Les méthodes et techniques de conception d'une solution intègrent son évaluation soit implicitement, sans technique prévue à cet effet, soit explicitement, grâce à une technique donnée. Toujours est-il qu'elles se basent sur un processus itératif. En effet, la maquette et le prototype sont des outils qui servent à stimuler les idées et à les raffiner à travers les itérations de conception jusqu'à une spécification qui corresponde aux besoins et caractéristiques des utilisateurs. Enfin, cette catégorie de méthodes fait appel à une activité de simulation plus ou moins sophistiquée qui permet aux utilisateurs de se projeter dans la situation future (Béguin, 2004a; Daniellou, 2004) et d'affiner la solution à concevoir. Elles peuvent mettre en œuvre une activité de maquettage ou de prototypage, généralement combinée à l'élaboration de scénarios (Carroll, 2000).

La maquette peut se réduire à une simple projection de transparents jusqu'à prendre la forme d'une maquette 3D manipulable, mais il est généralement impératif qu'elle soit amenée sur le lieu de travail de l'utilisateur (ex. écrans d'une interface projetés sur une boîte en carton faisant office d'ordinateur ; Ehn & Kyng, 1991). Ce dernier simule pas à pas le travail habituellement accompli à l'aide de la maquette (Muller, 1997) et le concepteur est censé lui indiquer les possibilités et limites du système. A chaque étape, les informations nécessaires à l'utilisateur peuvent être identifiées afin que la maquette soit modifiée en conséquence. Partant d'une maquette très simple, le processus itératif peut aboutir à un prototype électronique.

Plus axée encore vers une activité de conception coopérative, la méthode *PICTIVE*¹⁰ (Muller *et al.*, 1995) propose aux utilisateurs, encadrés d'un ou de plusieurs acteurs de l'entreprise¹¹, de réaliser un prototype qui représente l'apparence et les principes dynamiques de l'interface. Pour ce faire, ils utilisent des papiers, couleurs et matériels pré-imprimés représentant des éléments d'interfaces conformes à des guides de style ou relatifs au domaine spécifique (ex. icônes de l'interface existante). Tout en simulant des pans de leur activité avec ces prototypes-papiers, les utilisateurs modifient leurs agencements à volonté. Ces simulations se font sur la base de scénarios abordant les questions du « comment » et du « quand » (Darses & Reuzeau, 2004) qui permettent de décomposer la tâche principale en

¹⁰ Abréviation de Plastic Interface for Collaborative Technology Initiatives through Video Exploration.

¹¹ Ils peuvent être des professionnels de la conception, de l'ergonomie, du marketing, des formateurs, des rédacteurs techniques mais aussi des clients ou consommateurs des produits finis.

tâches de niveau inférieur ou de séquences d'action (Bisseret *et al.*, 1999). Les actions et moyens de réalisation de ces actions sont ainsi évalués. Les simulations peuvent aussi être filmées en vue d'être montrées à des managers ou concepteurs.

Dans le chapitre des techniques utilisant scénarios et prototypes, citons aussi la méthode *CISP*¹² (Madsen & Aiken, 1993) basée sur le principe du *storyboard prototyping*. Une petite équipe d'utilisateurs et de concepteurs réalise un story-board sous la forme d'un prototype dynamique (enchaînement d'écrans) à l'aide d'un outil de programmation approprié. L'utilisateur réalise la tâche requise sur le story-board puis toute l'équipe l'évalue en rejouant la séquence d'utilisation enregistrée grâce à l'outil logiciel. Il s'agit de répéter cela jusqu'à obtenir le story-board satisfaisant et donc l'interface adéquate. L'étude « terrain » de Beagley (1996) constitue un bon exemple d'utilisation de cette méthode appliquée au développement d'un système informatique militaire de repérage.

2.2.3. Méthodes et techniques d'évaluation participative

Certaines méthodes portent moins sur la génération d'une solution de conception et davantage sur l'évaluation d'une première forme existante de solution de conception. Quelques unes de ces techniques sont présentées ci-dessous. Elles sont parfois basées sur l'élaboration de scénarios comme les précédentes.

Par exemple, la technique *interface theatre* (Muller, 1997) est empruntée au monde du théâtre. Les membres de l'équipe de conception se créent des costumes et des accessoires représentant des boîtes de dialogue, curseur et autres éléments d'interface à taille humaine et démarrent en jouant le début d'un scénario qu'ils ont imaginé. Les acteurs tiennent ensuite compte des propositions d'« entrées-clavier » ou d'« actions-curseur » d'un public constitué d'utilisateurs, tout en se comportant conformément à des « scripts orientés objets » qui correspondent au fonctionnement établi de l'interface¹³. Le scénario ainsi déroulé est ensuite critiqué par ce public qui contribue à modifier l'apparence et les actions de l'interface. Les acteurs rejouent les « nouvelles interfaces » jusqu'à ce qu'il n'y ait plus de critiques.

Dans la technique plus classique d'évaluation coopérative (*cooperative evaluation* ; Wright & Monk, 1991)¹⁴, l'évaluateur, de préférence un concepteur, et un seul utilisateur explorent ensemble le prototype (déjà existant) et aboutissent à son analyse critique. La

caractéristique véritablement coopérative de cette technique tient au fait que l'utilisateur n'est pas seulement le sujet observé mais est aussi co-évaluateur. Il peut être amené à répondre à des questions du « pourquoi » qui permettent de préciser ses tâches à un niveau général (leur

¹² Abréviation de Cooperative Interactive of Storyboard Prototyping.

¹³ Une première solution de conception est donc déjà élaborée.

¹⁴ Cette technique peut aussi être utilisée dans l'évaluation par *think aloud*, décrite au paragraphe 2.4.2.2.

identification, leur objectif et leurs conditions; Bisseret *et al.*, 1999) et aussi à des questions du type « qu'est-ce que le système fait ou doit faire si... ? ».

On peut également évoquer les méthodes d'évaluation basées sur l'inspection d'une interface par des experts en ergonomie, voire par des concepteurs. Soit, comme les *pluralistic walkthroughs* (Bias, 1991), elles sont conçues pour faire participer plusieurs utilisateurs au travail d'inspection d'une équipe d'évaluateurs pluridisciplinaire (développeurs, architectes du système, rédacteurs techniques, etc.). Soit, elles sont aménagées à partir d'une technique individuelle, telle la *participatory heuristic evaluation* (Desurvire, 1994; Muller, 1997) où un ou plusieurs évaluateurs cherchent à détecter les problèmes de l'interface en prenant les points de vue de différents acteurs (utilisateur, ergonomiste, concepteur et autres acteurs). Cette inspection se fait sur la base d'un ensemble choisi d'heuristiques.

Après avoir présenté les méthodes et techniques « traditionnelles » de la conception participative, nous examinons le *focus group* tel qu'appliqué en ergonomie qui présente des similitudes évidentes avec ces dernières. Un seul titre suffit d'ailleurs à la justification du paragraphe suivant dans le grand ensemble « conception participative » de ce chapitre : *Beyond Focus Groups ? The use of Group Discussion Methods in Participatory Design ?* (Langford *et al.*, 2003).

2.2.4. Le focus group et la conception participative

Nous exposons d'abord ce qu'est un *focus group*, abordons ensuite les points de divergence concernant l'intégration du *focus group* dans le registre de la conception participative et enfin, présentons les usages du *focus group* qui s'apparentent à de la conception participative.

2.2.4.1. Qu'est-ce qu'un focus group

Le *focus group* est, en substance, une discussion de groupe modérée par un animateur qui doit centrer (*focus*) les discussions sur les informations recherchées et même les resserrer au fur et à mesure de l'évolution des échanges pour approfondir les points pertinents. Le groupe est généralement constitué de 5 à 10-12 participants. Il s'agit là de trouver le meilleur compromis entre trop peu d'interactions dans le groupe et trop de discussions en aparté entre quelques participants (Krueger & Casey, 2000; Merton *et al.*, 1956). Il est recommandé (Caplan, 1990; Krueger & Casey, 2000; Maguire, 2003; Merton *et al.*, 1956; Morgan &

Krueger, 1993) que chaque groupe présente une certaine homogénéité en termes d'âge, de milieux socioculturels et d'expérience des participants par rapport au produit ou au thème étudié. L'objectif est souvent d'obtenir une large variété de points de vue, ce qui requière de faire appel à différentes populations. Il est alors nécessaire de former un groupe pour chaque nouvelle population de sujets étudiée.

D'autre part, la durée d'une séance varie d'une à plusieurs heures. La séance peut avantageusement se prolonger jusqu'à quatre heures afin de ne pas l'interrompre durant la phase où le groupe est « mature » et disposé à révéler encore beaucoup d'informations pertinentes (Mariampolski, 1991). En effet, différentes phases peuvent être distinguées dans la formation d'un groupe (Gordon & Langmaid, 1988; Mucchielli, 1996)¹⁵. Les participants prennent d'abord connaissance des uns et des autres lors d'une phase d'échauffement¹⁶. S'ensuit généralement une phase d'affrontement car ils découvrent qu'ils ont en face d'eux des personnes aux opinions différentes. En vertu de l'effet « boule de neige » (Stewart & Shamdasani, 1990), l'émission d'une première opinion suscite toujours une série de réponses et, bien souvent, des points de vue opposés. Les participants sont ensuite amenés à expliciter leurs différents vécus. Ils parviennent alors à une phase de « politesse » où s'installe une véritable ambiance d'acceptation de l'autre. Enfin, le groupe débouche sur la phase de maturité ou de résolution où les participants, s'étant fait une idée générale de « qui pense quoi », osent davantage se parler franchement, tout en continuant à exposer leurs opinions sans s'affronter. Si ce déroulement a bien eu lieu¹⁷, ils ne cessent généralement plus d'ajouter des récits et de partager des expériences communes, sous l'effet de la stimulation (Stewart & Shamdasani, 1990). En dernier lieu, la phase de deuil permet de demander aux participants s'ils ont eu l'occasion d'exprimer tout ce qu'ils voulaient.

Pour l'animateur, il s'agit d'encourager, voire de tempérer pour rassurer les participants et recueillir *toutes* les opinions, lors des premières phases. Quand le groupe a atteint son niveau de maturité, il doit, en revanche, questionner pour creuser les opinions et arguments de chacun. Il doit notamment amener les participants à faire référence au vécu de leurs expériences (Merton *et al.*, 1956; Morgan, 1997) qui permet plus sûrement d'accéder aux motivations des participants. Ces dernières constituent des « forces inconscientes » qui

¹⁵ Mucchielli en distingue trois tandis que Gordon et Longmaid en distinguent cinq, lesquelles sont superposables au découpage de Mucchielli.

¹⁶ Des exercices d'échauffement (empruntés aux techniques du théâtre) ou des discussions entre deux participants peuvent être prévus lors de cette phase.

¹⁷ Diverses causes de blocage peuvent faire que le groupe ne parvient pas à ce stade (Gordon & Longmaid, 1988 ; Mucchielli, 1996).

déterminent les attitudes qui sont elles-mêmes des structures stabilisées et génératrices des opinions dont elles constituent les principes organisateurs (Mucchielli, 1996).

Finalement, l'animateur doit être formé à cette technique. Il ne s'agit pas nécessairement d'un spécialiste de l'animation (comme dans le cas d'une étude sous-traitée) ; il est souvent préférable que ce soit un professionnel du domaine étudié¹⁸ qui ait une bonne connaissance du thème à renseigner ou du produit à concevoir (Langford & McDonagh, 2003a; Morgan & Krueger, 1993). Il doit notamment avoir développé des astuces pour « maîtriser » les « leaders » qui cherchent à dominer la discussion, les « experts » qui peuvent être intimidants parce qu'ils savent tout, les « touristes » qui parlent pour ne rien dire et enfin pour encourager les plus réservés (Krueger & Casey, 2000; Merton *et al.*, 1956). L'animateur doit également être non directif sur le contenu (ex. pas d'expression d'opinion personnelle) mais directif quant à la forme, c'est-à-dire notamment maintenir le thème central de discussion et élucider ce qui est dit. Cela permet d'assurer la spontanéité des communications et la capacité créatrice d'idées nouvelles dans l'affrontement des situations en contrôlant « *les forces de blocage, de freinage, d'écran ou de modelage a priori des idées* » (Mucchielli, 1996, p. 42).

2.2.4.2. Le *focus group*, un outil pour une démarche participative ?

L'effet synergique des interactions du *focus group* permet d'obtenir une compréhension fine des opinions, des attitudes et des motivations à partir des récits des usages et expériences vécues des utilisateurs. Le *focus group* est donc généralement employé au début du processus de conception (a) pour connaître les besoins et caractéristiques d'utilisateurs représentatifs, (b) pour obtenir une première sélection de différents concepts de départ et (c) pour apprécier le niveau d'acceptabilité d'un concept (Baccino *et al.*, 2005; Maguire, 2003; Nielsen, 1993; Rubin, 1994; Van Vianen *et al.*, 1996). Il peut aussi servir à apprécier le niveau d'acceptabilité d'un produit déjà implanté sur le marché (phase de « feed-back », ex. Bastien & Scapin, 2004).

Le *focus group* pourrait être rapproché des différents *workshops* cités précédemment, dans la mesure où il rassemble des utilisateurs pour juger une situation ou un produit. Même si certains (Baccino *et al.*, 2005; Rubin, 1994; Van Vianen *et al.*, 1996) font une distinction entre les techniques de la conception participative et le *focus group* (surtout dans le domaine de la conception de produits informatiques et grand public), une explication plus intéressante nous semble possible, en considérant les divers domaines d'application de l'ergonomie

(situations de travail, du quotidien et même de relation de service). Il n'y aurait, en fait, pas de terminologie claire pour désigner le *focus group* (Langford & McDonagh, 2003a). Certains exemples, dans la littérature de ces différents domaines d'application, intitulent la méthode « groupes de discussions », « groupes de travail », *design decisions groups* (DDG) ou encore *user workshops*, alors qu'ils incorporent des aspects de cette technique. Cela est dû au fait que la notoriété du terme *focus group* s'est étendue depuis plus récemment et au fait que le concept de départ a été adapté et renommé par les ergonomes. De même que plusieurs auteurs (Bruseberg & McDonagh-Philip, 2001; Langford & McDonagh, 2003a), nous prenons le parti d'intituler *focus group* toute méthode incluant un niveau conséquent de discussions de groupe sans exclure d'autres activités collectives.

Dans ce qui suit, nous exposons les emplois du *focus group* dans deux domaines d'application de l'ergonomie qui présentent de nombreux points communs avec la conception participative, sans relever nécessairement d'une démarche au plus fort degré participatif.

Dans le domaine de la santé et de la sécurité au travail

Le *focus group* est une méthodologie établie dans le domaine de la recherche sur la santé et la sécurité au travail. Par exemple, des *focus groups* ont été menés pour étudier les effets

¹⁸ Il peut s'agir d'un chercheur (en sciences sociales) ou d'un concepteur.

méconnus de la prise de psychotropes¹⁹ sur les performances au travail²⁰ en termes de vigilance, mémoire et donc sécurité (Haslam, 2003). Cette étude a abouti à une série de recommandations qui concernent, d'une part, la prévention de l'anxiété et de la dépression au travail et, d'autre part, la prise en charge sanitaire des personnes qui souffrent de ces troubles. Des raisons qui expliquent les accidents dans l'industrie du bâtiment ont également pu être fournies grâce à une étude par *focus groups* (Haslam, 2003).

Dans la conception de produits/systèmes

- La génération de nouveaux concepts :

Les DDG (Langford & McDonagh, 2003b; Langford *et al.*, 2003), qui pourraient être nommés *focus groups*, constituent une méthode flexible qui permet de générer des concepts de produits/systèmes. Ils sont constitués d'un assemblage d'activités de discussion et de créativité en groupe pour analyser des problèmes en vue de concevoir (et implicitement d'évaluer) une solution plus ou moins aboutie.

Il peut s'agir, par exemple, de visionner la projection d'un film d'une activité donnée afin d'en tirer une analyse collective. Des exercices de *word map* (élicitation de mots portant sur un thème donné qui sont ensuite reliés les uns aux autres par association logique²¹), de *round-robin questionnaire* (où les participants complètent à tour de rôle un premier énoncé non achevé²²) ou de dessin (pour générer et partager des idées) permettent aussi de stimuler les idées, l'analyse critique et la recherche d'améliorations. Par ailleurs, des maquettes 2D ou 3D (plans de bâtiments, kits d'éléments-velcro associables, carton, terre glaise, etc.) assorties d'éventuels jeux de rôle peuvent aider les participants à travailler l'agencement d'un espace et son équipement, l'information affichée sur un écran ou un panneau, les dispositifs d'actionnement, le confort d'un poste de travail ou encore la forme d'un objet. Enfin, la « prise en main » de produits, éventuellement essayés dans un environnement simulé en grandeur nature, peut aider à l'évaluation de concepts existants et à la spécification de nouveaux (Langford *et al.*, 2003)²³.

¹⁹ Médicaments qui agissent sur le psychisme et donc sur les émotions et le comportement.

²⁰ En particulier concernant les professions les plus exposées en cas de difficultés de santé mentale (activité de soin, éducation, industries minières et manufacturières).

²¹ Cette technique est proche de la technique d'élicitation associative abordée au paragraphe 2.6.2.2.

²² De même, cette technique est très proche de la technique du *brainwriting*, abordée dans la description de la technique d'élicitation associative au paragraphe 2.6.2.2.

²³ Chacun de ces exercices sont aussi décrits en détail (indépendamment des DDG) dans Langford et McDonagh (2003b).

Une autre méthode de *focus group* permet la génération de concepts de produits (Cooper & Baber, 2003) : les participants y sont invités à imaginer des scénarios de leurs usages d'un produit du quotidien (ex. portefeuille) pour se projeter ensuite dans l'évaluation d'un prototype d'un produit du futur (ex. portefeuille électronique). Ils sont amenés à imaginer comment les caractéristiques du prototype pourraient accentuer les avantages et diminuer les inconvénients identifiés dans les scénarios de leur quotidien. Cette technique, très proche des *scenario-based techniques* de la conception participative, permet de dépasser les difficultés qu'ont les utilisateurs à imaginer concrètement leur activité avec un produit encore inexistant (Cooper & Baber, 2003).

- L'évaluation de prototypes et la génération de critères d'acceptabilité :

Toujours dans le domaine de la conception, le *focus group* peut aussi être un outil d'évaluation de prototypes (en le combinant à d'autres méthodes) et de génération de critères d'acceptabilité.

En effet, il offre l'avantage de faire participer les utilisateurs pour qu'ils partagent et explicitent leurs expériences après des sessions individuelles de test d'un prototype (Atyeo *et al.*, 1996; Caplan, 1990; Macaulay, 1996; Maguire, 2003; O'Donnell *et al.*, 1991; Sullivan, 1991). Il a été démontré que le *focus group* est un instrument de mesure valide : les tâches difficiles identifiées lors de tests utilisateurs, sont significativement davantage discutées que les tâches moins difficiles (O'Donnell *et al.*, 1991). Les interactions de groupe s'avèrent être un véritable moteur pour faire émerger ce que les utilisateurs ont à l'esprit au moment de l'utilisation du prototype. Ainsi, en plus d'une évaluation quantitative, les raisons explicatives aux problèmes rencontrés et d'éventuelles solutions qui y remédieraient peuvent être obtenues (O'Donnell *et al.*, 1991).

Une activité d'élaboration de critères d'acceptabilité peut aussi être menée en *focus group*. Cette technique peut être appliquée pour l'évaluation de plusieurs concepts (Caplan, 1990) et permettre de choisir entre différentes directions de conception (Langford & McDonagh, 2003b). Nous verrons au paragraphe 2.6.2.2. que des techniques similaires sont déjà utilisées depuis les années 1950 dans l'évaluation sensorielle de produits de consommation (alimentaires, cosmétiques).

Après avoir détaillé la plupart des méthodes de conception collective et/ou participative, nous explorons d'autres techniques utilisées en ergonomie pour la conception de produits/systèmes. Nous distinguons un ensemble de techniques faisant intervenir

l'utilisateur, d'une part, et des techniques d'évaluation sans l'utilisateur (ex. modèles formels, inspections d'interface par des experts ; paragraphe 2.5.), d'autre part. Les techniques faisant intervenir l'utilisateur peuvent être basées soit sur le recueil de données expérimentales (ex. tests utilisateurs ; paragraphe 2.3.), soit sur le recueil de données verbales (ex. protocoles verbaux ; paragraphe 2.4.)²⁴, soit sur le recueil de données subjectives (ex. évaluation subjective sensorielle, paragraphe 2.6. ; enquête par questionnaire sur l'utilisabilité et la satisfaction ; paragraphe 2.7.). Le paragraphe 2.6. porte aussi sur l'évaluation subjective sensorielle de produits de consommation, même si nous débordons là le champ de l'ergonomie. L'ensemble de toutes ces techniques porte sur le recueil d'informations concernant l'interaction de l'utilisateur/consommateur présent ou non avec le produit afin de les implémenter dans la conception.

Dans ce qui suit, nous abordons l'évaluation avec l'utilisateur, basée sur le recueil de données expérimentales.

2.3. LES DONNEES EXPERIMENTALES DANS L'EVALUATION

Le test utilisateur est une technique dont les débuts remontent à ceux de la conception participative, puisqu'en 1994, Desurvire (1994) affirme que l'évaluation empirique en laboratoire est la méthode de prédilection des ergonomes depuis au moins deux décennies (au moins concernant le développement de systèmes informatiques). Pourtant, dans les années 1980, les méthodes de laboratoire ou tests de performance n'ont pas nécessairement bonne réputation, peut-être en raison de la façon de recruter les participants ou de l'influence du personnel encadrant sur l'évaluation (Macleod, 1996). Ces évaluations ont aussi tendance à étudier des aspects très spécifiques et à tenir trop peu compte de l'environnement de travail, de la nature réelle des tâches des personnes concernées et de leurs caractéristiques, notamment socioprofessionnelles (Bannon & Bodker, 1991). D'autre part, les mesures de performance et résultats statistiques impressionnent peu les dirigeants des entreprises ou arrivent trop tard. Dans les années 1990, les laboratoires de test ont toutefois vu leur popularité s'améliorer, qu'il s'agisse des laboratoires du type « studio » avec équipement d'enregistrement audiovisuel ou des laboratoires ambulants qui peuvent être installés sur le lieu de travail (Macleod, 1996).

²⁴ Le plan proposé ici suit aussi une logique chronologique et par discipline. Ainsi, l'évaluation sans l'utilisateur est abordée après l'évaluation avec l'utilisateur, basée sur le recueil de données expérimentales ou verbales et avant l'évaluation avec l'utilisateur, subjective et/ou par questionnaire.

2.3.1. L'expérimentation contrôlée

La méthodologie suivie pour conduire un test trouve ses origines dans la psychologie expérimentale. Il s'agit de formuler une hypothèse et de la tester en isolant et manipulant des variables sous conditions contrôlées. L'hypothèse est définie en termes de variables indépendantes (celles que l'on manipule) et de variables dépendantes (celles qui sont mesurées). Elle postule qu'une variation de la variable indépendante engendrera une modification de la variable dépendante et constitue une prédiction du résultat expérimental (Baccino *et al.*, 2005; Dix *et al.*, 1993). Par exemple, on peut supposer une amélioration de la vitesse de recherche alors que le nombre d'items dans un menu diminue et que des commandes claviers sont utilisées plutôt que la « souris ». Les variables indépendantes peuvent aussi être le style de l'interface, le niveau de l'aide, le type d'icônes (etc.) tandis que les variables dépendantes sont des mesures dites de performance, voire des mesures subjectives comme la préférence de l'utilisateur.

Ces dernières sont relatives au concept de satisfaction tandis que les mesures de performance sont relatives aux concepts d'efficacité et d'efficience, tous trois définissant l'utilisabilité selon la norme ISO (ISO, 1998). L'efficacité est généralement évaluée par le nombre d'erreurs et de buts atteints et l'efficience par le temps de réalisation de la tâche. La méthodologie du projet ESPRIT MUSiC²⁵ (Macleod, 1996; Macleod *et al.*, 1997), qui se veut plus proche d'une réalité tangible, mesure également la qualité et la quantité de la performance de l'utilisateur, c'est-à-dire jusqu'à quel point les buts atteints sont complets et corrects.

Le choix des sujets est également primordial dans une expérimentation contrôlée. L'échantillon des sujets doit être représentatif de la population à étudier en termes d'âge, de niveau d'éducation, d'expérience professionnelle ; le mieux étant qu'il s'agisse d'utilisateurs réels (Dix *et al.*, 1993). La taille de l'échantillon doit aussi être valide sur un plan statistique. Nous revenons là-dessus plus loin. Enfin, il s'agit d'être attentif à divers éléments impactant les conditions expérimentales : l'ordre de passation, l'effet d'apprentissage dans les échantillons appariés, l'homogénéité des sujets entre le groupe « contrôle » et les autres groupes de sujets et entre les différents échantillons indépendants (Baccino *et al.*, 2005).

²⁵ Abréviation de Measuring Usability of Systems in Context.

2.3.2. Le test utilisateur

Dans l'industrie, pour des raisons liées à des contraintes temporelles et budgétaires, cette approche formelle n'est généralement pas suivie. Les tests utilisateurs sont donc réalisés sans tenir compte de toutes les exigences scientifiques relatives aux méthodologies expérimentales (Rubin, 1994), même s'ils poursuivent des objectifs similaires à leurs acolytes plus formels, à savoir la mesure de l'efficacité, de l'efficience et de la satisfaction²⁶.

2.3.2.1. Quel nombre de sujets faut-il prévoir ?

Un compromis récurrent et largement débattu au sein de la communauté ergonomique concerne le nombre de sujets nécessaires pour détecter un pourcentage d'erreurs jugé suffisant. La question posée est : comment obtenir des résultats valides en investissant le moins de ressources ? A partir de quel nombre d'utilisateurs, les nouvelles informations fournies ne valent-elles plus la peine d'être recueillies d'un point de vue productivité ?

Concrètement, les praticiens n'ont souvent accès qu'à cinq ou six opérateurs du terrain (Thomas, 1996) ou même qu'à quatre ou cinq (Davis, 1996) pour réaliser des tests ou observations/entretiens informels et ils obtiennent néanmoins des résultats pertinents pour améliorer la conception des produits/systèmes. Sur un plan scientifique, une formule mathématique désormais célèbre (Nielsen & Landauer, 1993) prédit que cinq tests utilisateurs (formels) permettent de détecter 85% des problèmes d'utilisabilité et que 15 tests utilisateurs en détectent la totalité²⁷. Cependant, cette formule est vivement critiquée (Woolrych & Cockton, 2001) car elle ne tient pas compte des spécificités de la situation d'évaluation (complexité et du produit à évaluer et de la tâche d'utilisation) et des différences interindividuelles des sujets (capacité à détecter un problème et variation du niveau de sévérité attribué aux problèmes). Les travaux de Nielsen et Landauer font courir le risque de passer à côté de problèmes sévères (Baccino *et al.*, 2005). D'autres travaux ont révélé que 10 et 20 utilisateurs permettent de détecter respectivement un minimum de 80% et 95% des problèmes d'utilisabilité (Faulkner, 2003). Un échantillon de 15 à 20 utilisateurs semble donc recommandé du point de vue de la fiabilité statistique (Baccino *et al.*, 2005), même si c'est rarement pratiqué en milieu industriel.

²⁶ Et même si parfois des mesures temporelles ne peuvent pas être recueillies (absence d'outils de recueil et d'analyse de « fichiers log ») (Thomas, 1996).

²⁷ Nielsen (1993) a pourtant décrit les limitations de cette recommandation. Selon Faulkner (2003), à l'époque, il était plus efficace pour les ergonomes praticiens de retenir et de prôner une formule simple dans un contexte où l'ergonomie était encore imparfaitement acceptée.

2.3.2.2. Les données complémentaires au test utilisateur

Tout un ensemble de données peut compléter celles obtenues lors du test utilisateur : des mesures d'indicateurs comportementaux, des données verbales et des données concernant la satisfaction de l'utilisateur. Les mouvements oculaires permettent, par exemple, de faire certaines inférences sur le processus cognitif de recherche d'informations (Baccino *et al.*, 2005). Les données verbales peuvent aussi avantageusement compléter les données comportementales en labellisant des informations traitées par les sujets et insaisissables avec les seuls observables comportementaux. Ainsi, des techniques de protocoles verbaux (explicitées au paragraphe 2.4.2.2.) peuvent accompagner le test utilisateur. Enfin, des entretiens et questionnaires peuvent être prévus à la fin du test afin d'évaluer la satisfaction de l'utilisateur ou d'autres éléments subjectifs (détaillés respectivement aux paragraphes 2.4.2.2. et 2.7.1.2.).

2.3.2.3. Différents types de tests utilisateurs

De même que les méthodes de la conception participative, le test utilisateur peut s'appliquer très en amont dans la conception. Il s'agit alors du test exploratoire. On peut aussi distinguer le test d'évaluation, le test de validation et les tests comparatifs (Rubin, 1994).

Comme son nom l'indique, le premier type de test est plus une exploration d'un concept ou d'un aspect du futur produit, représenté sous la forme d'une maquette-papier ou d'un prototype sommaire, alors que les spécifications sont éventuellement encore en discussion. Dans le cas d'une application logicielle, les fonctionnalités majeures peuvent être représentées graphiquement sans être fonctionnelles (prototype horizontal) ou bien une ou quelques fonctionnalités seulement modélisées de façon fonctionnelle (prototype vertical) (Nielsen, 1993). Par ailleurs, la simulation peut être assurée par des techniques allant de l'évaluation papier-crayon²⁸ (Bevan, 2002; Maguire & Bevan, 2002; Rubin, 1994) à la technique dite du magicien d'Oz²⁹. Il peut être demandé au sujet d'accomplir une tâche représentative ou de simplement parcourir la modélisation donnée, s'il n'existe pas encore de tâches suffisamment définies. Ce type de test informel ressemble davantage à une évaluation coopérative où l'expérimentateur sollicite l'utilisateur pour comprendre la nature des difficultés qui se présentent à lui et les raisons pour lesquelles il procède de la façon dont il le fait. Cela permet de savoir comment l'utilisateur conçoit le produit, si les fonctionnalités de

²⁸ Où l'enchaînement des écrans-papier est assuré par l'expérimentateur en fonction des choix de l'utilisateur (appelé aussi *paper prototyping*).

²⁹ Où la réponse de l'ordinateur est simulée par un compère caché de l'utilisateur (le magicien).

départ lui semblent utiles, si les diverses opérations et la navigation sont intuitives, quel type d'aides et d'informations sont nécessaires pour utiliser certaines fonctionnalités, etc.

Si le but du test exploratoire est de travailler à partir du squelette du produit, avec le test d'évaluation, il s'agit de travailler sur le corps et la chair du produit. Il faut cette fois-ci détecter les erreurs d'utilisabilité du produit. L'utilisateur accomplit des tâches représentatives avec un prototype plus avancé, l'expérimentateur l'observe plus que de l'interroger et des mesures de performances sont recueillies (accompagnées, si les moyens le permettent, du recueil de données comportementales).

Dans le test de validation, il s'agit de vérifier si le produit opérationnel est acceptable avant sa mise sur le marché. C'est souvent seulement à cette occasion que tous les composants du produit (manuel d'utilisation, accessoires de branchement, etc.) peuvent être testés ensemble. Même si les enjeux ne permettent pas de renoncer au lancement immédiat au profit d'améliorations supplémentaires, le chef de projet pourra toujours préparer des stratégies de « récupération » : lancer des projets de développement de « patches »³⁰, instruire le service clients ou la « hotline » des réponses à dispenser aux clients, etc. Ce type de test a surtout un intérêt au sein d'une entreprise qui cherche à établir des normes d'utilisabilité pour une même famille de produits (Rubin, 1994). Par exemple, s'il a été identifié qu'une procédure d'installation d'un « pack » logiciel prend 15 minutes, les produits suivants devront être au moins à la hauteur de cette performance.

Enfin, les tests comparatifs peuvent être pratiqués à n'importe quel stade. Il peut s'agir de comparer simplement deux fonctionnalités, deux concepts de produits ou deux produits opérationnels concurrents.

Malgré ces diverses précautions, une situation de test présente toujours une certaine artificialité. Les participants ne sont jamais parfaitement représentatifs de la future population d'utilisateurs et les mesures de probabilités garantissent seulement que les résultats ne sont pas le fruit du hasard mais non que le produit fonctionnera, c'est-à-dire sera effectivement utilisé efficacement (Rubin, 1994). Nous savons, en effet, que la conception se poursuit dans l'usage et qu'il s'agirait de prendre pour objet central de recherche les processus d'utilisation et non l'artefact lui-même (Bannon & Bodker, 1991; Rabardel, 1995). Des usages hors normes voire informels peuvent notamment se développer. Le sujet peut donc attribuer une signification imprévue à un artefact, en regard de la conception telle qu'arrêtée par les

³⁰ Petits programmes téléchargeables qui permettent la réparation des « bugs » informatiques.

concepteurs. Enfin, réaliser des tests sur des aspects très spécifiques du produit/système et/ou en situation de laboratoire contribue à négliger la signification donnée au produit/système dans l'activité en situation naturelle. Dans ce qui suit, nous étudions cette notion de signification ou de subjectivité à travers l'histoire de l'utilisation des verbalisations.

2.4. LES DONNEES VERBALES DANS L'EVALUATION

Nous décrivons ici la lente habilitation des verbalisations dans la psychologie en général et dans l'évaluation des produits en particulier. Nous exposons aussi plusieurs techniques d'évaluation avec l'utilisateur basées plus spécifiquement sur le recueil de données verbales (même si certaines des précédentes techniques abordées jusque là peuvent également recueillir ce type de données en accompagnement).

Les verbalisations sont une production, le plus souvent orale mais parfois aussi écrite, d'énoncés en langue naturelle qui ne sont pas spontanés mais provoqués à des fins de recherche (Caverni, 1988). Les données verbales sont utilisées depuis les débuts de la psychologie, constituée en tant que science, pour comprendre l'expérience humaine. Les psychologues ont, en effet, des difficultés à reposer sur les seules observations externes pour rendre compte des processus mentaux. Cependant, il faut attendre les années 1980 pour que les verbalisations soient véritablement reconnues comme des données valides.

2.4.1. L'histoire de l'utilisation des verbalisations et de la subjectivité en psychologie

2.4.1.1. Le rejet des données verbales en psychologie

La « première forme » de données verbales est relative à l'activité d'introspection. La psychologie trouve ses origines dans la philosophie qui, depuis la Grèce Antique, a cherché à comprendre la nature et le contenu des phénomènes qui influencent l'expérience humaine. Certains philosophes ont cherché à étudier l'esprit humain en se basant sur l'introspection, c'est-à-dire l'étude de leurs propres pensées.

Cette méthode d'observation directe est aussi celle des débuts de la psychologie. Toutefois, étant donné le caractère subjectif de ces observations individuelles, la psychologie cherche à développer des méthodes plus scientifiques, aussi par analogie aux sciences naturelles (biologie, physique, etc.). Le mouvement du « structuralisme » va écarter tout ce

qui est de l'ordre d'inférences et de généralisations à partir de l'expérience consciente et chercher à recueillir uniquement les faits de la conscience (sa description directe). Cette méthode, dite de l'introspection classique, entraîne les sujets à relater une forme trapézoïdale rouge plutôt qu'un livre rouge lorsque ce stimulus leur est présenté (Scane, 1987). Ce mouvement présente l'inconvénient majeur de réduire la conscience à des composants sensoriels élémentaires et de supprimer ce qui est connaissance et expérience constituées liées aux objets.

Les *behavioristes* et plus particulièrement Watson critiquent cette approche mais n'apportent pas pour autant d'améliorations, en considérant que la psychologie peut se passer de toute donnée d'introspection et de toute construction mentale. Ils remplacent l'étude des événements mentaux par l'étude du seul comportement. Leur focalisation sur la performance manifeste a renforcé leur suspicion à l'égard des données verbales. Pourtant, selon Ericsson et Simon (1984, 1993), ils ne considèrent pas les données de l'introspection comme nécessairement fausses mais comme non nécessaires. Watson ferait même une distinction entre la méthode de l'introspection classique qu'il réfute, le fait de demander à un sujet de s'exprimer oralement et le fait de lui faire dire les seules pensées qui lui viennent à l'esprit en lien avec la réalisation de sa tâche (la technique du *think aloud* qu'il ne réfute pas). Néanmoins, Watson considère aussi la pensée comme un discours intérieur et le discours comme une habitude laryngienne ! Les personnes privées de larynx, qui pourtant pouvaient penser, lui posaient problème (Scane, 1987).

Toujours est-il que l'ère du *behaviorisme* engendre un déclin dans l'utilisation des données verbales et jusque dans les années 1970, plusieurs critiques sont formulées à l'égard des données verbales (ex. Nisbett & Wilson, 1977). Ericsson et Simon y répondent, au milieu des années 1980, avec leur ouvrage *Protocol Analysis : Verbal Reports as Data* (Ericsson & Simon, 1984, 1993). Ils montrent en quoi et à quelles conditions les verbalisations doivent être considérées aussi valides que les mesures comportementales. Nous revenons là-dessus après avoir décrit le mouvement convergent de plusieurs disciplines des sciences humaines vers l'acceptation de la subjectivité.

2.4.1.2. L'évolution vers l'acceptation de la subjectivité

Un mouvement convergent de plusieurs disciplines

Ericsson et Simon (1984, 1993) notent aussi, sans doute grâce à leur ouvrage, que les années 1980 ainsi que le début des années 1990 sont marquées par une nette augmentation de

l'utilisation des données verbales pour l'étude des processus cognitifs en psychologie, en sciences de l'éducation et en sciences cognitives. Cette époque correspond aussi à celle où des insuffisances sont dénoncées dans la psychologie des HCI. Selon *The Kittle House Manifesto* (Carroll, 1991), l'usage des méthodes de laboratoire comme l'application des théories du traitement de l'information produisent des résultats peu convaincants. Ces approches sont remises en question. Il y a besoin de prendre en compte la véritable nature de la tâche dans l'environnement familier du sujet considéré en rapport avec sa culture et son histoire, et ainsi de replacer au centre les questions de la signification pour lui (Bannon & Bodker, 1991). Par ailleurs, à la même époque, la recherche qualitative commence à être acceptée comme apportant des données aussi valides que la recherche quantitative, en sciences sociales appliquées comme en marketing (cf. § 2.1.2.)³¹. Ces recherches qualitatives sont basées sur le recueil de la signification pour le sujet, via les verbalisations, même si les consignes de verbalisation ne sont pas aussi formalisées que celles définies par Ericsson et Simon.

Les années 1980-1990 marquent ainsi un tournant dans différentes disciplines des sciences humaines. De plus, la réhabilitation des verbalisations a un lien évident avec celle de la subjectivité.

La psychologie élève la subjectivité au rang de concept explicatif

Dans ces différentes disciplines, l'époque du « tout expérimental » et du « tout quantitatif » cède peu à peu la place à des approches où l'on accepte et veut mieux connaître le vécu subjectif de l'individu. En psychologie plus particulièrement, le *behaviorisme*, puis le cognitivisme sont dénoncés pour s'être fourvoyés, en tant que sciences positivistes, avec leurs « idéaux de réductionnisme, d'explication causale et de prédiction » (Bruner, 1991, p. 15). L'explication causale oblige à comprendre par avance le phénomène à observer et contraint d'artificialiser ce qui est étudié au point qu'il est difficile d'y reconnaître une représentation de la vie humaine. Au contraire, l'explication plausible repose sur la signification construite par chaque individu dans son environnement historique, social et culturel. La signification est publique, partagée par les individus d'un même groupe ethnique et social, maintenue grâce à des procédures de négociation et d'interprétation, et donne à l'individu une stabilité dans sa vie sociale.

³¹ Les verbalisations sont d'ailleurs utilisées de façon informelle depuis les années 1930 en sciences sociales appliquées comme en psychologie.

La psychologie doit donc s'intéresser, non pas au comportement qui correspond au seul « faire » qui exclut le « dire », mais à l'action fondée sur l'intentionnalité et plus précisément à l'action située, c'est-à-dire interprétable dans un ensemble culturel et en fonction des interactions réciproques des intentions des participants. En réhabilitant cette notion d'interprétation (et du récit), Bruner (1991, p. 30) « élève la subjectivité au rang de concept explicatif ». Il affirme aussi que ce phénomène est déjà à l'œuvre depuis plus longtemps dans des disciplines-sœurs des sciences humaines comme l'anthropologie, la philosophie, la linguistique, etc.

L'importance du langage mis en évidence par la psychanalyse et l'école de la psychologie russe

On ne saurait clore ce développement sans faire allusion à la psychanalyse, à la psychologie génétique et surtout à l'école de la psychologie russe.

Dès la fin du 19^{ième} siècle, après un court passage par l'hypnose, Freud développe des thérapies psychanalytiques basées sur le principe selon lequel la parole et la libre association des idées joue un rôle primordial dans les processus de prise de conscience et de libération des affects. On notera, par exemple, que dans la psychanalyse du petit Hans (Freud, 1954, 1999), des verbatim sont rapportés intégralement par le père qui prend des notes des dialogues qu'il a avec son fils, alors que les enregistrements audio n'existent pas encore à cette époque.

Piaget s'appuie également, entre autres, sur les réponses verbales de ses sujets pour étudier le développement mental chez l'enfant, mais il ne s'intéresse pas véritablement au rôle du langage dans la pensée. C'est Vygotski qui s'attache à examiner le rôle du langage (et notamment son caractère d'emblée socialisé) dans la formation de la pensée et du développement de l'individu (Vergnaud, 1999). Selon l'hypothèse dégagée par Friedrich (1999), l'intérêt porté au langage amène même Vygotski à s'éloigner de ses collaborateurs et de Léontiev en particulier, lequel fonde le mouvement des Kharkoviens et s'intéresse davantage à l'activité et à l'étude de la communication à travers l'activité. Vygotski, probablement en raison de ses intérêts linguistiques (entre autres), a une véritable « prise de conscience de l'importance des faits langagiers pour la psychologie » et se met à étudier de plus près les théories linguistiques contemporaines pour proposer « une articulation intéressante entre une analyse de la langue et la problématique qui est propre à la psychologie, à savoir celle des affects, des émotions, de la pensée » (Friedrich, 1999, p. 183). Vygotski meurt en 1934. En outre, dans les années 1930, avec l'arrivée de Staline au pouvoir,

la psychotechnique disparaît quasiment. Elle ne se relève que dans les années 1960 avec le dégel initié par Khrouchev (Nosulenko & Samoylenko, 1999a).

Les principaux psychologues de cette période sont Léontiev et Lomov qui développent les théories de l'activité (individuelle et commune) et surtout, la psychologie de l'ingénierie ou approche systémique de la psychologie (dirigée par Lomov). « *Selon cette approche, la cognition humaine ne peut être étudiée indépendamment des processus de communication qui lui sont associés. La communication humaine et l'activité sont, de ce point de vue, les deux formes différentes et interdépendantes d'un même être social* » (Nosulenko & Samoylenko, 1997)³².

Dans la psychologie « occidentale », ce n'est que dans les années 1980 que des écrits comme ceux de Vygotski, de Léontiev et de Lomov sont revisités, alors que la psychologie est atteinte d'une seconde révolution, celle du « contextualisme » (nécessité d'étudier la cognition en situation) après la « révolution cognitive » des années 1950, toutes deux initiées par Bruner.

Ayant ainsi replacé l'émergence de l'intérêt porté à la subjectivité et aux données verbales dans l'histoire de la psychologie, nous exposons maintenant leur application contemporaine dans le domaine de l'élicitation de connaissances et de l'évaluation de produits/systèmes.

2.4.2. Pertinence et utilisation des verbalisations pour l'élicitation de connaissances et l'évaluation de produits/systèmes

Des travaux d'autres chercheurs russes cherchent à vérifier la pertinence des verbalisations pour l'analyse de processus psychiques. Une étude d'Artemyeva (1980 ; citée par Nosulenko & Samoylenko, 1997, 1999a) consistait notamment à demander à des sujets de verbaliser les propriétés de plusieurs objets qui leur étaient présentés. A partir des descriptions verbales, d'autres sujets étaient à même de reconnaître quel objet était associé à chaque description. Cette démarche inversée contribue à asseoir la validité des verbalisations. De même, des études russes (Urvanstev, 1979, 1994 ; citées par Nosulenko & Samoylenko, 1997, 1999a) mais aussi américaines (Loftus, 1972; Russo & Rosen, 1975) ont montré une correspondance entre la trajectoire des mouvements oculaires ou le nombre de fixations oculaires et les objets qui sont verbalisés.

³² Nosulenko et Samoylenko (1997) font référence à Lomov (1984) sans pour autant le citer littéralement.

Ericsson et Simon (1980, 1984, 1993) replacent également le problème de validité des protocoles verbaux en référence « *d'une part à un modèle du fonctionnement cognitif pour la tâche considérée (tâche principale, activité cible), d'autre part à un modèle, complémentaire, de la production verbale (tâche annexe)* » et « *s'attachent à énoncer les principes généraux de l'un et de l'autre modèle* » (Caverni, 1988, p.267). Il s'agit donc d'examiner en quoi et jusqu'à quel degré les verbalisations reflètent l'activité principale (cognitive, perceptive).

2.4.2.1. Les consignes de verbalisation

Pour Ericsson et Simon (Ericsson & Simon, 1980, 1984, 1993), la qualité des verbalisations et des inférences qu'elles permettent de réaliser quant aux processus étudiés, dépend en grande partie de la consigne de verbalisation. Si l'on demande aux sujets de verbaliser uniquement les pensées qui entrent dans leur champ d'attention en lien avec la tâche réalisée, le processus cognitif (ou aussi la séquence de pensées) demeure inchangé. Dans la technique qu'ils proposent, il s'agit de chercher à n'obtenir que les informations, déjà codées verbalement ou facilement transposables verbalement, qui traversent l'esprit du sujet au moment où il réalise une tâche donnée ou immédiatement après le déroulement d'une action très courte (0,5 à 10 secondes maximum). Il s'agit donc de faire verbaliser les informations présentes dans la mémoire à court terme. Il ne faut demander aucune explication qui pourrait engendrer des processus d'interprétation et de reconstruction a posteriori. Ces verbalisations doivent être très proches de celles du sujet qui se « parle tout haut » en réalisant, par exemple, une opération de calcul mental.

Des exercices préalables d'entraînement à ce type de verbalisations peuvent être prévus (Ericsson & Simon, 1984, 1993; Karat, 1997), à la suite de quoi, une consigne du type « dites tout haut ce que vous vous dites à vous-mêmes lorsque vous réalisez la tâche » est donnée au sujet. Enfin, seules des relances du type « continuez à parler » (au lieu de la demande plus sociale « dites-moi à quoi vous pensez ») seront employées. L'expérimentateur doit résister à l'envie d'obtenir des verbalisations plus complètes ou plus cohérentes, sous peine de recueillir des données non valides³³. D'autres données (ex. mouvements oculaires) aident à compléter leur interprétation, si nécessaire, car les verbalisations recueillies par *think aloud* permettent néanmoins un recensement complet de l'information qui est *effectivement* prise en compte pendant la réalisation de la tâche considérée.

³³ Les verbalisations recueillies par *think aloud* contiennent des séquences interrompues avec différentes pensées sans liens explicites entre elles, en comparaison à des verbalisations du type « communications sociales » qui suscitent justifications et rationalisations.

Cependant, deux autres types de consignes sont aussi employés : « Expliquez comment vous faites pour exécuter la tâche » pour verbaliser la procédure ou « Expliquez ce que vous faites et pourquoi vous le faites » pour justifier la procédure (Bisseret *et al.*, 1999). Ces deux consignes engendrent davantage de distorsion. Des mécanismes de prise de conscience vont s'opérer lors de la verbalisation de la procédure. Généralement, seules les représentations des buts et des résultats sont présentes chez le sujet lors de l'exécution de l'action. Il peut ensuite s'opérer une prise de conscience des moyens de réalisation, puis des raisons qui font que ces moyens sont appropriés et efficaces (Piaget, 1974). En verbalisant et en re-simulant mentalement la procédure, le sujet va s'orienter vers la reconstitution de la structure de la procédure (Bisseret *et al.*, 1999). Avec la troisième consigne, lorsque le sujet se force à ou est forcé de justifier la procédure, il atteint le niveau de conscience le plus élevé quant à son activité (Hoc & Leplat, 1983). Le sujet va alors analyser la coordination générale des différentes actions dans son activité et rechercher les raisons qui ont contribué aux résultats obtenus. A ce niveau, ses verbalisations sont le fruit de constructions a posteriori, extérieures aux mécanismes d'exécution d'une opération ou d'élaboration d'une procédure.

Néanmoins, l'utilisation de cette dernière consigne peut être justifiée par la volonté d'accéder aux motivations conscientes du sujet et/ou de l'utiliser dans un contexte d'apprentissage car elle a un effet d'amélioration de la performance des sujets (Bisseret *et al.*, 1999). Ericsson et Simon (1984, 1993) privilégient la consigne de penser tout haut car il s'intéressent à la suite des opérations mentales et non aux causes explicites pour le sujet de ces opérations.

Enfin, selon Nosulenko et Samoylenko (1997, 1999a), un certain type de protocole expérimental est une condition nécessaire à l'obtention de verbalisations pertinentes de l'activité psychique (perceptive et cognitive). Il s'agit d'intégrer une opération de comparaison dans l'activité de verbalisation lors de l'évaluation d'objets.

Diverses méthodes se sont développées sur la base de ces consignes de verbalisation. Nous les détaillons dans ce qui suit.

2.4.2.2. Les méthodes utilisant les verbalisations

Protocoles verbaux, évaluations par think aloud et techniques d'autoconfrontation

A l'origine, la technique des protocoles verbaux est utilisée en psychologie pour étudier des tâches de reconnaissance de stimuli visuels, d'apprentissage de concepts complexes, de

résolution de problème, etc. Cette technique a ensuite été appliquée en psychologie cognitive et en psychologie ergonomique pour étudier les processus de jugements et de prise de décisions des experts, leur performance, leurs connaissances et leurs procédures. L'étude de ces derniers, en particulier, doit servir à la conception de systèmes experts. En matière de conception, il s'agit de faire verbaliser tout haut l'expert ou tout utilisateur lorsqu'il réalise sa tâche en situation naturelle (Baccino *et al.*, 2005). Cela peut toutefois aussi être réalisé en situation de simulation de l'activité (Bisseret *et al.*, 1999; Daniellou, 2004). Cela permet, en général en phase de conception initiale, de formaliser les tâches réalisées, d'identifier les fonctions nécessaires au futur système et d'organiser son interface en fonction des procédures mises en œuvre par les utilisateurs (Bastien & Scapin, 2004).

L'évaluation par *think aloud* est un dérivé de la technique des protocoles verbaux et se déroule en faisant utiliser une maquette ou un prototype plus ou moins finalisé (Baccino *et al.*, 2005). Dans ce cas, il y a donc interaction avec une représentation matérielle du futur produit, contrairement aux protocoles verbaux où il n'y en n'a pas, particulièrement dans le cas d'une situation à informatiser. L'évaluation par *think aloud* est souvent utilisée consécutivement à des protocoles verbaux qui ont servi à l'élaboration du prototype. Elle permet d'affiner la définition des fonctionnalités en cours de conception. Elle peut aussi servir à estimer le niveau d'utilisabilité du produit fini (Baccino *et al.*, 2005), même si elle est plus appropriée au début du processus itératif, car il s'agit d'une technique assez lourde et produisant des données qualitatives (Karat, 1997).

Des techniques de verbalisations consécutives assistées par les traces (autoconfrontation) peuvent aussi être utilisées. Il s'agit de recueillir les verbalisations du sujet alors qu'il lit la transcription des verbalisations, écoute l'enregistrement audio ou visualise l'enregistrement vidéo de sa propre activité (Bisseret *et al.*, 1999), voire de celle d'autres sujets réalisant la même activité (Mollo & Falzon, 2004). Enfin, il peut être demandé au sujet de verbaliser à partir de chacune de ses opérations, en déroulant l'enregistrement visuel pas à pas (Hoc & Leplat, 1983). Ces verbalisations consécutives semblent aussi performantes que des verbalisations simultanées sans en présenter les inconvénients (ralentissement de l'élaboration et de l'automatisation de la procédure). Pourtant, elles constituent un décalage supplémentaire d'avec l'activité réelle.

Certaines classifications (Baccino *et al.*, 2005; Dix *et al.*, 1993)³⁴ catégorisent les techniques décrites ci-dessus dans les méthodes observationnelles, en distinction des

méthodes analytiques (inspection de l'utilisabilité), des méthodes basées sur un modèle de la tâche, des méthodes expérimentales et des méthodes de questionnement (questionnaire, entretien). Nous proposons de regrouper ces techniques dans le grand ensemble « données verbales », étant donné leur problématique originelle liée à la nécessité de démontrer la validité de ces données en face de celle des données comportementales. De plus, nous nous intéressons particulièrement aux techniques qui utilisent les verbalisations dans le cadre de cette thèse. Enfin, les entretiens doivent aussi être considérés dans cette catégorie.

Les entretiens

Les entretiens proviennent de la psychologie clinique ou sociale. Ils sont en quelque sorte une variante simplifiée des protocoles verbaux. Les verbalisations n'étant alors ni simultanées ni consécutives (au sens défini par Ericsson et Simon), elles peuvent être moins valides.

Néanmoins, les entretiens permettent de recueillir des informations sur la population des futurs utilisateurs et d'orienter les choix de conception (Baccino *et al.*, 2005; Maguire & Bevan, 2002), quand ils sont utilisés très tôt dans le processus de conception. Ils permettent aussi de connaître l'expérience d'utilisation, une fois l'usage du produit/système élaboré et consolidé par les utilisateurs (Bastien & Scapin, 2004; Dix *et al.*, 1993) et d'apprendre comment le produit est approprié (intégré à des schèmes préexistants), voire instrumentalisé par les utilisateurs (Rabardel, 1995). Bien que rares (Rubin, 1994), ces études apportent des connaissances pour concevoir de nouveaux produits/systèmes à partir de l'usage réel (ex. fonctionnalités effectivement utilisées ou non, usages non prévus au départ, etc.).

Deux formes d'entretien sont distingués : (a) les entretiens non-directifs, dits aussi libres ou approfondis, où l'intervieweur propose un thème sans trop le spécifier et n'intervient que pour relancer le sujet qui peut s'éloigner des questions posées, (b) les entretiens semi directifs ou structurés, où l'intervieweur cherche à recentrer la discussion sur les différents thèmes qu'il peut aborder dans n'importe quel ordre (Baccino *et al.*, 2005; Ghiglione & Matalon, 1998)³⁵. De plus, les entretiens peuvent être menés (a) seuls, (b) après un test utilisateur ou (c) au poste de travail pour que l'opérateur et l'évaluateur puissent se référer aisément au système utilisé et discuter des points critiques (Baccino *et al.*, 2005; Mayhew, 1999).

³⁴ Baccino *et al.* (2005) réalisent un découpage encore légèrement différent de celui de Dix *et al.* (1993).

³⁵ En sociologie, trois formes d'entretiens sont distinguées avec, en plus des deux déjà citées, l'entretien directif qui est très similaire à un questionnaire administré en situation d'entretien.

La dernière situation peut aussi aider l'utilisateur à se souvenir de son usage habituel et des difficultés rencontrées. Un autre type d'entretien facilitant cette activité de mémoire sans l'aide de « traces » est l'entretien d'explicitation conçu par Vermersch (1994, 2000).

L'entretien d'explicitation

L'entretien d'explicitation cherche à rendre conscientisées les connaissances « implicites » contenues dans l'action. Comme nous l'avons déjà évoqué, les travaux de Piaget montrent qu'il y a possibilité de réussite sans la conceptualisation et qu'un processus de prise de conscience des moyens et raisons de la réussite peut s'opérer. La mise au jour de connaissances non conscientisées va passer par les différentes étapes de la prise de conscience. L'entretien d'explicitation se base là-dessus ; il constitue « *une prise de conscience provoquée* » (Vermersch, 2000, p. 84). Concrètement, le sujet identifie une séquence d'activité précise sur laquelle portera l'entretien et doit être placé dans une « *position de parole incarnée* ». Il faut l'amener à ce qu'il « *soit en train de "revivre", d'évoquer de manière concrète la situation de référence* » (Vermersch, 2000, p. 57). Cet entretien permet d'accéder au détail des actions effectives dans une tâche donnée, de leur enchaînement et de leur articulation avec les prises d'informations. En ce sens, l'entretien d'explicitation ressemble grandement aux techniques des protocoles verbaux, à ceci près qu'il ne s'agit ni de verbalisations simultanées, ni de verbalisations consécutives. De plus, contrairement aux protocoles verbaux, l'entretien d'explicitation permet d'accéder à certains éléments du comportement routinier (Amalberti & Hoc, 1998)³⁶.

A l'origine, cet entretien est mis au point pour décrire ce qui fait l'essentiel de l'expertise d'un professionnel (dans le secteur pédagogique, en particulier) et pour l'aider à se perfectionner, mais l'usage de cet entretien s'est beaucoup développé depuis dans le domaine de l'évaluation ergonomique (ex. Leulier *et al.*, 1999).

Les verbalisations recueillies selon les différentes techniques décrites ci-dessus peuvent toutes faire l'objet d'une analyse de contenu.

³⁶ En effet, il est recommandé de s'appuyer sur des moyens complémentaires (mouvements oculaires, activité de dessin, techniques d'autoconfrontation) (Amalberti & Hoc, 1998; Bisseret *et al.*, 1999; Leplat & Hoc, 1981), quand on soupçonne une forte composante routinière ou sensori-motrice dans les activités étudiées (dont on sait qu'elles ne peuvent pas être verbalisées) ou pour des informations difficiles à verbaliser comme les représentations spatiales.

2.4.2.3. Modalités de recueil des données verbales et analyse de contenu

Concernant tout d'abord les modalités de recueil des données, il faut savoir que la possibilité d'enregistrer les données verbales sur bande sonore a contribué à obtenir des données dites « dures » (Ericsson & Simon, 1984, 1993). La prise de notes qui se pratiquait auparavant (et encore maintenant au niveau industriel quand les contraintes de temps sont fortes) incorporait de l'interprétation du seul fait que l'expérimentateur filtrait ce qui était dit car il ne pouvait pas tout noter.

Cependant, les données ne sont « dures » que lorsque leur recueil *et* leur interprétation sont suffisamment solides (Bisseret *et al.*, 1999; Ericsson & Simon, 1984, 1993). Les inférences réalisées à partir des données verbales et donc les règles d'analyse doivent être claires et systématiques. Le traitement consiste en une segmentation des données et un codage des segments obtenus. Segmenter le protocole verbal peut se faire par épisodes (ex. d'une pause à une autre dans le discours ou correspondant à un enchaînement d'actions pour l'atteinte d'un sous-but) ou par thèmes (un segment se détache du précédent dès lors qu'un nouveau thème est abordé). Généralement, coder suppose d'établir des thèmes ou catégories d'analyse relatifs à ce que l'on cherche à renseigner. Le modèle de l'activité ou de la tâche, si on en dispose, peut aider à cela (Caverni, 1988). Le codage s'appuie ainsi sur une taxonomie plus ou moins exhaustive des opérations réalisables dans le contexte de la tâche (ex. objets sur lesquels le sujet peut agir et actions possibles sur ces objets). L'idéal est de disposer à la fois du protocole verbal et du protocole d'exécution (ex. enregistrement vidéo de l'activité) ce qui facilite leur codage mutuel.

Une méthode « inter juges » peut aussi être employée (Bisseret *et al.*, 1999; Caverni, 1988). Elle consiste à croiser les traitements de plusieurs experts. Plus prosaïquement, il est aussi possible de relever les idées qui ont émergées dans une série d'entretiens à partir de notes et de compter le nombre de sujets ayant évoqué chaque idée (Mayhew, 1999). En effet, dans le milieu industriel, il est fréquent de ne travailler qu'à partir des notes, malgré les réelles imprécisions engendrées.

Après avoir abordé les techniques basées sur le recueil de données expérimentales ou verbales avec l'utilisateur, nous examinons l'évaluation sans l'utilisateur. Cette nouvelle forme d'évaluation se développe parallèlement à l'intensification de l'utilisation des verbalisations, pour des raisons économiques.

2.5. L'ÉVALUATION SANS L'UTILISATEUR

Si sur le plan scientifique, les années 1990 sont celles de la réhabilitation des données verbales et subjectives, sur le plan social, ce sont celles de l'accroissement des contraintes de temps et des exigences en termes de qualité.

Avec la reconstruction d'après guerre, il s'agit de satisfaire une demande conséquente mais moins exigeante que celle des années 1990³⁷. En effet, après la tendance « technologique » des années 1970³⁸, à partir des années 1980, il y a une demande du public orientée vers des objets faciles à utiliser. Les technologies informatiques (Spérando, 1985) et, plus tard, celles des télécommunications (Atyeo & Green, 1994) se multiplient et se « popularisent ». La nécessité de concevoir des produits/services simples, adaptés au grand public se fait sentir.

Par ailleurs, la compétition industrielle s'accroît et les industriels prennent conscience de la valeur ajoutée que constitue l'ergonomie dans leurs produits (Jordan *et al.*, 1996a; Jordan *et al.*, 1996b). Néanmoins, compétition rime aussi avec concevoir plus vite et moins cher (Desurvire, 1994). Les temps de mise au point et d'évaluation du produit se trouvent réduits (Gaillard, 1997; Wright & Monk, 1991). Les ergonomes développent de nouvelles techniques basées sur l'évaluation par un expert, cherchant à remplacer les techniques plus coûteuses faisant intervenir l'utilisateur (test utilisateur, observation, entretien, questionnaire, *focus group*, les « fichiers log » ; De Vries *et al.*, 1996; Desurvire, 1994). Ils élaborent les modèles formels (ex. Card *et al.*, 1983), le *cognitive walkthrough* (Polson *et al.*, 1992), l'inspection basée sur des heuristiques (Nielsen, 1994b; Nielsen & Molich, 1990), sur des critères ergonomiques (Bastien & Scapin, 1993; Leulier *et al.*, 1998), puis plus tard sur les normes ISO (ISO, 1996).

Avec les modèles formels, non seulement l'utilisateur, mais aussi les éléments du système sont représentés par un modèle (Whitefield *et al.*, 1991). Avec les méthodes d'inspection (*cognitive walkthrough*, évaluation heuristique, critères ergonomiques, etc.), seul l'utilisateur est représenté par l'évaluateur. L'objectif recherché est que ce dernier s'emploie à utiliser le produit/système comme l'utilisateur le ferait, en l'explorant librement ou en suivant

³⁷ Même si les années 1950 et 1960 font face à une compétition croissante avec le boom industriel d'après-guerre et stimulent le développement du marketing (surtout aux États-Unis).

³⁸ Une approche technocentrique domine en raison des résistances des concepteurs à la prise en compte de l'utilisateur et de la culture de l'époque (ce qui est complexe à utiliser signifie que c'est « technologique » et donc quelque chose de valable en soi).

des scénarios de tâches étudiés pour être représentatifs de l'activité des utilisateurs (Baccino *et al.*, 2005; Karat, 1994).

2.5.1. Les modèles formels

Le développement des modèles formels démarre dès les années 1980. Ils reposent sur une analyse de ce qu'un individu doit faire pour utiliser un système donné, à partir d'une représentation théorique de cet utilisateur. Les modèles les plus connus, GOMS³⁹ (Card *et al.*, 1983) et CCT⁴⁰ (Kieras & Polson, 1985), de même que le *cognitive walkthrough*, reposent sur un même principe d'une « structure de but » (*goal structure*) où le but final est décomposable en sous-buts ou actions élémentaires pour la réalisation d'une tâche (Polson *et al.*, 1992). Les modèles GOMS et CCT sont généralement décrits sous la forme de diagramme d'actions ou implémentés informatiquement et permettent de simuler la séquence ordonnée des sous-buts (Baccino *et al.*, 2005; Polson *et al.*, 1992; Shneiderman, 1998). Ces modèles cherchent à modéliser les routines employées par des experts pour réaliser une tâche en situation IHO⁴¹ et à fournir des prédictions concernant les performances des utilisateurs (temps de réalisation, temps de formation nécessaire).

Les modèles formels sont destinés à produire des spécifications avant que toute solution de conception n'existe (Whitefield *et al.*, 1991) et/ou à aider à choisir parmi différentes alternatives (Dix *et al.*, 1993). Cependant, leur application ne saurait se substituer à une activité de prototypage et de recueil de données empiriques (Bannon & Bodker, 1991). En effet, la représentation théorique de l'utilisateur sur laquelle se basent ces approches relève du cognitivisme et de la théorie du traitement de l'information qui considère la cognition humaine comme un système de manipulation de symboles, abstraits et non situés. Un écart est généré entre ces modélisations et l'activité en situation naturelle, puisqu'elles reposent sur ce que l'utilisateur devrait faire et non sur ce qu'il fait réellement. Ces approches, de même que les approches similaires de *task analysis* (ex. Stanton & Baber, 1996a), sont donc incapables de rendre compte des tâches telles qu'elles se déroulent en situation naturelle (Bannon & Bodker, 1991).

³⁹ Abréviation de Goals Operators Methods and Selection.

⁴⁰ Abréviation de Cognitive Complexity Theory.

⁴¹ Abréviation de Interface Homme-Ordinateur.

2.5.2. Les méthodes d'inspection de l'utilisabilité

Le *cognitive walkthrough* et les méthodes d'inspection à partir d'heuristiques ou de critères ergonomiques constituent une évolution par rapport à ces modèles, même si cette évolution demeure insuffisante. Nous débouchons sur ces critiques, après les avoir explicitées.

Par rapport aux précédents modèles, le *cognitive walkthrough* est destiné à évaluer une solution de conception (maquette, prototype) et cherche à dépasser l'étude des processus routiniers (Polson *et al.*, 1992). Le but du *cognitive walkthrough* est de déterminer si la connaissance générale de l'utilisateur et les éléments fournis à l'interface sont suffisants pour que l'utilisateur se construise la « structure de but » nécessaire à l'élaboration de la séquence d'actions requises pour accomplir sa tâche. Une étude préalable des caractéristiques des futurs utilisateurs et des spécifications de l'interface sert à établir des scénarios de tâches comprenant cette séquence d'actions. L'évaluateur la simule ensuite sur l'interface et, à chaque étape, il doit vérifier si les buts de l'utilisateur et les actions nécessaires pour manipuler l'interface correspondent (« action » visible, distincte des autres, *feedback* du système compréhensible, etc. ; Baccino *et al.*, 2005).

De même que le *cognitive walkthrough*, à l'origine, les méthodes d'inspection utilisant les heuristiques ou les critères ergonomiques ont été développées pour les concepteurs, afin qu'ils puissent réaliser des évaluations rapides des interfaces qu'ils conçoivent. L'idée est qu'en analysant les interfaces sur la base de dimensions ergonomiques souhaitables, ils détectent les problèmes majeurs. Or, les *guidelines*, élaborés jusque là, constituent des listes abstraites et exhaustives, difficilement utilisables par les concepteurs (Bannon & Bodker, 1991; Gould & Lewis, 1985). Le peu d'études concernant la validité et l'utilisabilité des divers guides de recommandations ou normes (ISO, 1996; Marshall *et al.*, 1987; Ravden & Johnson, 1989; Smith & Mosier, 1986; Vanderdonck, 1995) est également dénoncé (Bastien *et al.*, 1998; Scapin & Bastien, 1997). Nielsen et Mollich, d'une part, Bastien et Scapin, d'autre part⁴², tentent d'ordonner ce chaos et de produire des listes de dimensions ergonomiques plus « digestes » et plus appropriées (valides, sensibles, fidèles, faciles à utiliser, etc.) pour les concepteurs.

Ces méthodes sont souvent utilisées dès les premières phases itératives de la conception (Baccino *et al.*, 2005; Karat, 1994) à partir d'une maquette-papier, d'un prototype plus

⁴² Chacun à leur façon, mais dont le résultat s'avère finalement très similaire.

élaboré ou parfois à partir du produit opérationnel (Baccino *et al.*, 2005). Il s'agit de décrire les problèmes d'utilisabilité, hiérarchisés en fonction de leur degré de sévérité, de même que les aspects positifs du produit/système.

2.5.3. Avantages et inconvénients des méthodes d'inspection de l'utilisabilité

Malgré les efforts réalisés, il s'avère que toutes les méthodes d'inspection tendent à requérir les compétences d'un ergonomiste. Le nombre élevé d'études (ex. Bastien & Scapin, 1992, 1995) qui comparent les performances d'évaluateurs novices/experts en ergonomie et/ou en informatique témoignent du souci de vérifier si les concepteurs sont susceptibles d'appliquer efficacement la méthode d'inspection à partir de dimensions ergonomiques. Or, il apparaît (a) que les concepteurs ont du mal à utiliser la méthode (Molich & Nielsen, 1990; Nielsen & Molich, 1990), (b) que les experts en ergonomie sont plus performants que les développeurs informaticiens (Desurvire, 1994), (c) que les experts en ergonomie sont plus performants que les novices même s'ils peuvent manifester une certaine résistance aux dimensions présentées car, à travers leur expérience, ils se sont construits leurs propres dimensions informelles (Bastien *et al.*, 1998), et enfin (d) que les professionnels possédant une double expertise (en ergonomie et en informatique) sont les plus performants par rapport aux autres groupes d'évaluateurs (Nielsen, 1992). De plus, les professionnels ayant des connaissances en sciences cognitives ont plus de facilité à utiliser le *cognitive walkthrough* (Polson *et al.*, 1992).

Néanmoins, ces méthodes peuvent présenter des intérêts adjacents. Par exemple, l'évaluation heuristique peut contribuer à faire comprendre concrètement au commanditaire les différents aspects que peut couvrir une démarche ergonomique, ce qui peut l'amener à valider l'intérêt d'un travail plus approfondi (Baccino *et al.*, 2005). Par ailleurs, les méthodes d'inspection peuvent présenter l'avantage de sensibiliser les concepteurs à la « conception centrée utilisateurs » (Bastien *et al.*, 1998; Scapin & Bastien, 1997). C'est d'autant plus vrai avec les *pluralistic walkthroughs* (Bias, 1991) conçus pour faire participer des professionnels de différentes compétences et profiter de la richesse de la mise en commun de leurs différents points de vue. Des séances de débriefing impliquant évaluateurs et concepteurs peuvent aussi être prévues après un *walkthrough* (Karat, 1994) ou une évaluation heuristique (Nielsen, 1994a).

Enfin, il s'avère que toutes ces méthodes d'inspection ne sont pas aussi performantes que les tests utilisateurs à détecter les problèmes d'utilisabilité (Desurvire, 1994; Karat, 1994). Par exemple, Macleod *et al.* (1997) affirment que les utilisateurs réalisent, quasiment à chaque fois, des choses inattendues lors de tests et rencontrent des problèmes que les évaluateurs auraient été incapables d'anticiper. La question de la validité des données recueillies avec les méthodes d'inspection et de celle des procédés de caractérisation de la sévérité des problèmes retenus reste posée (Mack & Nielsen, 1993). Enfin, la variabilité de jugement d'un évaluateur à un autre est aussi à soulever car elle dépend des compétences de ce dernier (Karat, 1994; Mack & Nielsen, 1993).

Comme beaucoup de méthodes, il est préconisé de les utiliser en combinaison et non en remplacement des données recueillies empiriquement. Cependant, le principe de ces approches est basé sur une simple représentation de l'utilisateur et engendrera nécessairement un écart d'avec l'activité réelle. La conception devrait plutôt trouver ses origines dans la prise en compte de la pratique développée par l'utilisateur en situation naturelle, avec ses dimensions socioculturelles (Bannon & Bodker, 1991). Dans le même ordre d'idées, les méthodes d'inspection passent à côté des questions de la satisfaction et d'autres aspects hédoniques et émotionnels qui interviennent dans l'interaction entre l'utilisateur et le produit. Nous explorons maintenant les techniques d'évaluation avec l'utilisateur/consommateur (qui peut cependant être entraîné à l'évaluation), basées sur le recueil de données subjectives, sensorielles.

2.6. L'ÉVALUATION SUBJECTIVE ET SENSORIELLE

Même si les contraintes budgétaires amènent une partie des industriels (et chercheurs) à s'orienter vers des méthodes d'évaluation sans l'utilisateur, le renouement avec la notion de subjectivité observé dans les années 1980 et 1990 poursuit son chemin. L'évaluation subjective (Les_Cahiers_du_Club_CRIN, 1997) ou sensorielle (SSHA, 1998; Urdapilleta, 2001a) s'est ainsi développée dans de nombreux domaines industriels, depuis l'alimentaire au cosmétique, en passant par les produits d'entretien ou encore l'automobile. Elle est utilisée pour intégrer le « facteur humain »⁴³ dans la conception de nouveaux produits (Les_Cahiers_du_Club_CRIN, 1997; Urban & Hauser, 1993) tout autant que pour évaluer leur qualité perçue, une fois consommés par le public. Un groupe de travail pluridisciplinaire a même été fondé. Les dernières connaissances constituées dans le champ de la logique floue, de l'analyse multi-critère et des statistiques ont été mises à profit pour développer des outils de mesure du subjectif (Les_Cahiers_du_Club_CRIN, 1997).

De même que dans le secteur des produits/services technologiques, les industriels des produits de consommation font face à une montée de la compétition et de l'exigence de la clientèle (Les_Cahiers_du_Club_CRIN, 1997; Parasuraman *et al.*, 1985). Toutefois, ils ne cherchent pas à réduire les ressources mises en œuvre pour le développement de leurs produits (ce qui est l'objectif poursuivi avec les méthodes d'évaluation sans l'utilisateur). Les entreprises ont recours à des panels de sujets formés à l'évaluation, alors que dans les années 1950, elles s'appuyaient seulement sur l'évaluation de leurs produits par quelques employés experts (Stone *et al.*, 1974). Cette discipline utilise néanmoins le sujet humain comme instrument de mesure aussi bien que des mesures instrumentales⁴⁴ et des mesures statistiques de consommation. De plus, le sujet peut être naïf mais, souvent, il s'agit d'un sujet amené à suivre une formation à l'évaluation plus ou moins longue, voire d'un expert dans certains domaines en particulier (ex. œnologue). Enfin, des observations comportementales (face à plusieurs produits, lequel choisit le sujet ?) sont préférées par certains chercheurs de cette discipline (Urdapilleta, 2001a).

⁴³ Au sens générique : prise en compte de l'humain dans la conception des produits alimentaires, cosmétiques, etc. aussi bien que des produits où l'ergonomie peut intervenir. Cela ne fait pas référence au mouvement des Human Factors.

⁴⁴ Description « objective » d'un produit obtenue à l'aide d'appareils de mesure. Par exemple, la mesure du pH d'une crème cosmétique.

2.6.1. L'histoire de l'évaluation sensorielle

2.6.1.1. Des débuts relatifs à la psychophysique

L'évaluation sensorielle connaît aussi une histoire assez ancienne. Les premiers à étudier les perceptions sont des psychophysiciens de la fin du XIX^{ième} et du début du XX^{ième} siècles. Fechner et Stevens (cités par Mac Leod *et al.*, 1998; Urdapilleta, 2001a) cherchent à produire une mesure dite objective des perceptions et à établir une loi universelle reliant intensité de la perception et stimulus. Partant du principe selon lequel, seule une différence entre sensations est perceptible par l'homme, Fechner a élaboré une loi qui porte son nom. Elle établit une relation logarithmique entre perception et stimulus : « *lorsque les intensités de la stimulation sont faibles, seul un faible changement dans la stimulation est nécessaire pour percevoir une différence, alors que lorsque les intensités de la stimulation sont fortes, un important changement dans la stimulation permet de percevoir la même différence* » (Urdapilleta, 2001a, p. XX). Stevens, quant à lui, part du principe que l'évaluation directe est possible et met au point une méthode d'estimation de la grandeur d'un stimulus à partir d'un premier stimulus de référence. Selon lui, un sujet est à même de dire si le stimulus présenté est deux fois plus fort ou trois fois moins que le stimulus de référence.

Fechner et Stevens ont donc permis de décrire les relations stimulus-réponse, ce qui s'apparente à une perspective behavioriste. Cela soulève à nouveau la question de quelle activité mentale se cache derrière les mesures psychophysiques.

2.6.1.2. La perception et l'activité mentale

Percevoir un objet ce n'est pas seulement réagir à un stimulus. L'activité électrophysiologique suscitée au niveau des organes sensoriels est transformée en information brute (image sensorielle) qui est progressivement filtrée, réduite, stabilisée et surtout contrastée à mesure qu'elle chemine vers les centres supérieurs du cerveau. La mémoire intervient à ce niveau. La perception donnée va être comparée aux références mémorisées et reconnue comme appartenant à une catégorie d'objets semblables (Mac Leod *et al.*, 1998). Ainsi, les réponses des sujets sont influencées par leurs représentations cognitives des référents objets des stimuli. Des stratégies décisionnelles sont aussi élaborées (relatives, par exemple, à la motivation des sujets) et influencent leurs réponses et jugements (Dacremont, 2001; Urdapilleta, 2001a).

Les travaux des psychologues de l'école russe contribuent également à mieux connaître l'activité mentale en lien avec la perception. Dès les années 1960 (et même avant malgré le nettoyage idéologique en URSS à partir des années 1930), des psychologues comme Rubinstein, Léontiev et Lomov mènent des travaux sur les concepts d'activité et de reflet psychique (Nosulenko & Samoylenko, 1999a). Leurs travaux contribuent à montrer en quoi le reflet psychique (reflet de la réalité objective telle que perçue par le sujet) est étroitement lié à l'activité ou interaction avec le monde dans laquelle est engagée le sujet. L'image psychique se forme donc sur la base de l'expérience de l'individu et dépend de ses besoins, motifs, buts, attitudes, émotions, etc. Ainsi, dans la perception (qui est un processus de création d'une image psychique), se manifestent les propriétés pertinentes des objets et les caractéristiques opérationnelles de l'activité. Autrement dit, le sujet ne perçoit pas seulement l'objet, mais aussi l'activité possible en rapport avec l'objet.

Lors d'une conférence de Nosulenko à l'Université de ParisVIII (pas de publication), ce dernier évoquait des tests perceptifs de bruits de portière de voiture lors desquels aucun sujet n'avait tout d'abord reconnu l'origine des bruits. Les sujets verbalisaient seulement leurs aspects forts, faibles, aigus, etc. Après précision de leur nature et présentation d'une photo d'une personne fermant la porte d'un véhicule, des jugements relatifs, non seulement aux propriétés acoustiques des stimuli, mais aussi aux actions possibles étaient obtenus dans les verbalisations des sujets. Dans le premier cas, seules les impressions détachées de toute expérience tangible (un peu comme dans la méthode de l'introspection classique) sont relatées tandis que dans le second cas, l'expérience et la connaissance du sujet se manifestent.

Enfin, dans ce phénomène, la dimension sociale doit aussi être considérée car la communication influence la perception et vice-versa (Lahlou *et al.*, 2002; Nosulenko & Samoylenko, 1999a). Les dimensions socioculturelles sont des éléments pris en compte par de nombreuses méthodes d'évaluation subjective, comme nous le verrons dans l'exposé qui suit.

Nous détaillons plusieurs méthodes d'évaluation qui prévoient la construction de dimensions subjectives : (a) les méthodes collectives et individuelles dans le domaine des produits alimentaires ; (b) la *repertory grid*, ses variantes et l'approche systémique de la psychologie appliquée à la qualité perçue, toutes appliquées en ergonomie. Enfin, nous abordons brièvement les évaluations selon des dimensions construites a priori.

2.6.2. Les méthodes d'évaluation intégrant l'élicitation de descripteurs ou d'attributs

Dans l'évaluation sensorielle/subjective, il s'agit non seulement de différencier l'intensité de stimuli, mais aussi de les qualifier, c'est-à-dire de décrire les dimensions pertinentes qui les caractérisent. Des techniques d'élaboration de descripteurs ont été développées (même s'il existe peu de mots pour décrire, par exemple, nos perceptions liées aux odeurs et aux timbres musicaux, en tous cas dans les langues indo-européennes ; Urdapilleta, 2001a, 2001b). Chaque stimulus ou produit est ensuite quantifié sur chacun des descripteurs et un profil sensoriel ou portrait verbal est ainsi établi.

2.6.2.1. Activités collectives et individuelles d'élaboration de descripteurs dans le domaine des produits alimentaires

Dans le domaine des produits alimentaires, de nombreuses méthodes ont été développées pour élaborer des listes de descripteurs. Le premier est Arthur D. Little (cité par Stone *et al.*, 1974; Urdapilleta, 2001a) qui, à la fin des années 1940, a mis au point une méthode basée sur des discussions de groupe (« méthode par consensus ») pour la qualification et la quantification des produits. A l'époque, cette méthode constitue une volonté de « s'affranchir » de la dépendance vis-à-vis des quelques employés experts de la dégustation. Des sujets sont invités à produire une liste commune de descripteurs à partir de l'essai d'échantillons de produits de référence. Ce travail est modéré et animé par un meneur de groupe. Les sujets quantifient ensuite les produits individuellement selon les termes retenus, puis les données sont rapportées à l'expérimentateur qui mène une seconde discussion de groupe afin de parvenir à un profil sensoriel unique et consensuel pour chaque produit sur la base des différentes données individuelles.

Cette méthode n'a eu de cesse de se transformer. Par exemple, la *Quantitative Descriptive Analysis* (Stone *et al.*, 1974) repose davantage sur une activité individuelle. La liste de descripteurs est rapidement obtenue à travers une discussion de groupe, puis les produits sont évalués individuellement et de façon répétée par les sujets, sur le long terme. Cette répétition permet de vérifier la constance des sujets, de poursuivre la formation de certains, voire d'en exclure. Les évaluations sont ensuite simplement moyennées.

Une autre méthode de profil conventionnel a été normalisée par l'AFNOR (Barthélémy, 1998; ISO, 1995; Urdapilleta, 2001b). La liste de descripteurs est obtenue par un travail individuel (élicitation d'attributs à partir de la dégustation d'échantillons) et de discussion en

groupe (prise de décisions collectives quant à l'élimination des synonymes, etc.). Cette liste est ensuite réduite par une série de traitements statistiques. Pour ce faire, les sujets évaluent les échantillons selon l'ensemble des descripteurs obtenus, puis l'expérimentateur va classer les descripteurs à partir de la moyenne géométrique de ces évaluations et éliminer ceux aux moyennes faibles. Un dernier calcul statistique (identification des descripteurs corrélés positivement – synonymes – ou négativement – antonymes) permet encore de réduire la précédente liste obtenue à une quinzaine de termes seulement. Enfin, les sujets sont entraînés à employer cette liste de descripteurs et notamment à s'accorder sur leur signification qui varie souvent d'un individu à un autre.

Tenant compte de ces variabilités interindividuelles, Williams et Langron (1984) mettent au point une méthode de *free choice profiling* où chaque sujet élabore et utilise sa propre liste de descripteurs. Chaque sujet peut évaluer les produits avec les mots qui lui paraissent intimement les plus à mêmes de représenter ses perceptions. Un outil mathématique permet ensuite de positionner les produits les uns par rapport aux autres selon les deux axes prépondérants pour le groupe de sujets (par analyse de la variance et en général pour différents thèmes, par exemple, pour l'arôme, la couleur ou la flaveur de vins).

Enfin, une nouvelle méthode (Urdapilleta, 2001b) est mise au point sur la base de la critique des précédentes. En effet, la méthode du profil conventionnel de l'AFNOR « *ne permet que peu d'échanges entre les sujets, alors que c'est par une confrontation des perceptions et des labels qui leur sont attachés que chacun peut affiner son opinion concernant les descripteurs* » (Urdapilleta, 2001b, p. 104). La méthode participative élaborée par Urdapilleta permet, au contraire, d'utiliser le groupe comme un opérateur statistique naturel (de même que la « méthode par consensus » d'Arthur D. Little). D'autre part, dans la méthode du profil avec libre choix de vocabulaire, le sujet produit et organise ses descripteurs sans aucun échange extérieur. Or, « *un travail de groupe a pour effet d'améliorer la finesse des descripteurs retenus en amenant chacun à énoncer les critères qu'il utilise* » (Urdapilleta, 2001b, p. 105).

Dans la nouvelle méthode, les sujets, par sous-groupe de trois, élicitent et s'accordent sur une liste de descripteurs. Chaque sous-groupe rapporte ensuite le fruit de ce pré-consensus à l'ensemble du groupe. Une discussion collective amène alors tous les sujets à une liste commune de descripteurs, où l'animateur s'assure que chaque descripteur correspond à la même définition pour l'ensemble des sujets. A nouveau par petits groupes, les sujets sont ensuite invités à produire une définition commune de chaque descripteur et à les associer à

une référence concrète. Une dernière discussion de groupe permet d'harmoniser les travaux de chaque sous-groupe.

2.6.2.2. Méthodes et techniques d'élicitation d'attributs en ergonomie

La repertory grid et techniques de catégorisation

La *repertory grid* (Kelly, 1955) est la technique d'élicitation d'attributs la plus ancienne et la plus connue. Kelly l'a mise au point pour étudier les *personal constructs* des individus, c'est-à-dire leurs conceptualisations personnelles des choses, des gens, des situations, etc. qui doivent expliquer, non seulement leurs comportements, mais aussi leurs émotions. Cependant, cette technique peut s'utiliser en se démarquant de la théorie de la personnalité conçue par Kelly (Bisseret *et al.*, 1999; Dillon & McKnight, 1990). Ainsi, à ses débuts, elle est beaucoup employée en psychologie clinique (Levy-Leboyer *et al.*, 1985), puis subit un coup de jouvence vers la fin des années 1980 et connaît, depuis, divers domaines d'applications : (a) en psychologie sociale et psychologie du travail pour étudier, par exemple, les orientations professionnelles ou le milieu socioprofessionnel d'un individu (Levy-Leboyer *et al.*, 1985), (b) en ergonomie, en vue de l'élicitation des connaissances des experts (Bisseret *et al.*, 1999; Dillon & McKnight, 1990) et aussi (c) de l'évaluation de type « benchmark » de la qualité globale⁴⁵ de produits/services (Baber, 1996; Hassenzahl & Trautmann, 2001) et enfin (d) en marketing, pour analyser les différences perçues entre marques et aider les stratégies du marketing (Gordon & Langmaid, 1988).

La *repertory grid* est, en fait, l'association de trois techniques élémentaires : celle de l'élicitation associative (détaillée plus loin), celle des triades et celle des échelles bipolaires (Bisseret *et al.*, 1999). La technique d'élicitation associative n'est pas toujours réalisée (Dillon & McKnight, 1990). Néanmoins, dans son principe, elle consiste à faire produire une liste d'objets, concrets ou symboliques, qui sont spécifiques au sujet et sur lesquels portera le travail ultérieur d'élaboration des attributs. Ces objets ou ceux choisis par le chercheur sont ensuite présentés par triades au sujet qui doit isoler celui qui lui semble différent des deux autres et expliciter le critère sous-jacent à cette différence. Les objets sont ensuite évalués selon ces critères soit après chaque critère énoncé, soit une fois l'ensemble de ceux « élicatables » obtenus. Ces derniers sont présentés au sujet avec leurs valeurs extrêmes inscrites aux deux bouts d'un segment (ex. prend du temps à lire – peut être lu rapidement).

Un ensemble de constructions-évaluations propres à chaque sujet est obtenu. Divers traitements statistiques peuvent être appliqués aux données généralement organisées sous la forme d'un tableau (*grid*) à double entrée (objets/critères d'évaluation) (Levy-Leboyer *et al.*, 1985), pour extraire des comparaisons inter-sujets, inter-objets, inter-attributs (Bisseret *et al.*, 1999). Cette technique permet donc d'obtenir, par exemple, un positionnement relatif de produits évalués en fonction des construits propres aux sujets.

Par ailleurs, l'application de la seule technique par triades ou de toute autre technique de catégorisation (Bisseret *et al.*, 1999) permet aussi de générer des attributs et de construire des échelles d'évaluation à items verbaux (Urdapilleta, 2001c).

La technique SHIRA⁴⁶

D'autre part, la *repertory grid* a inspiré des auteurs qui en ont tiré une autre technique d'évaluation subjective des produits ; *SHIRA* (Hassenzahl *et al.*, 2001a; Hassenzahl *et al.*, 2001b). Les auteurs cherchent à conserver, autant que faire se peut, le principe de travailler à partir des attributs pertinents pour le sujet⁴⁷, tout en optimisant la technique. En effet, la *repertory grid* requière du temps (même si elle est plus rapide que des entretiens approfondis), un professionnel formé à la technique et un minimum de quatre produits à comparer.

Ainsi, *SHIRA* est basée sur l'emploi d'un jeu considérable d'attributs relatifs à la qualité ergonomique (contrôlable, simple, etc.) et à la qualité hédonique (original, impressionnant, innovant, etc.). Le produit à évaluer est d'abord présenté au sujet, de préférence dans son contexte d'usage. Le sujet choisit un premier attribut important pour lui (ex. fiable) et doit ensuite relever toutes les caractéristiques du produit en rapport avec cet attribut (ex. besoin de pouvoir compter sur des commandes qui s'exécutent correctement). Enfin, il est invité à réfléchir aux solutions qui permettraient de répondre à chacune de ses attentes (ex. le système doit répéter les paramètres entrés). La technique fournit donc des résultats d'évaluation et des principes de solution de conception.

⁴⁵ Il s'agit là d'évaluer à la fois des aspects ergonomiques du produit (utilité, utilisabilité, etc.) mais aussi de faire émerger des dimensions très spécifiques au domaine de produit (ex. existence d'une fonction de décongélation ou non) et relatives aux caractéristiques socioculturelles et aux besoins des sujets.

⁴⁶ Abréviation de Structured Hierarchical Interviewing for Requirements Analysis.

⁴⁷ La technique cherche notamment à éviter le biais du différentiel sémantique qui est un ensemble de dimensions d'évaluation « pré-choisies » selon lesquelles le sujet doit donner une mesure du produit/service grâce à une échelle numérique (ex. de 1 à 5) ou à items verbaux (« très agréable », « assez agréable », « sans opinion », « peu agréable », « pas du tout »).

La technique d'élicitation associative

Empruntée au courant de l'associationnisme qui remonte aux études des philosophes grecs s'intéressant aux lois de l'association des idées (Scane, 1987), la technique d'élicitation associative repose sur le principe de faire réagir un sujet à un stimulus verbal (Bisseret *et al.*, 1999). La consigne est du type « décrivez » ou « écrivez rapidement tous les éléments qui vous viennent à l'esprit concernant... » le thème d'étude donné. Il s'agit de recueillir des mots, des expressions ou idées qui correspondent à une liste de problèmes relatifs à une activité donnée, aux thèmes que devrait traiter un manuel d'aide, aux critères d'évaluation d'une classe de produits, etc. Cette technique est très proche du *brainstorming* ou de l'équivalent sous forme écrite, le *brainwriting* (Warfield, 1976 ; cité par Boy, 1996a). Elle peut aussi être menée en passation collective. Chaque sujet écrit alors une liste d'éléments et, au bout d'un court laps de temps, passe sa feuille à son voisin qui réagit aux éléments qui lui sont soumis. Cela se poursuit jusqu'à ce que chaque feuille ait été vue par l'ensemble des sujets.

La technique d'élicitation associative constitue généralement une phase amont pour générer un corpus préalable à d'autres études. Par exemple, dans la *group elicitation method* (Boy, 1996a, 1996b), l'élicitation associative, en passation collective, est associée à des activités collectives et individuelles (catégorisation des éléments élicités en concepts généraux, hiérarchisation selon leur degré de priorité, discussion de groupe sur les résultats obtenus) pour mettre en évidence des problèmes ou avantages d'une situation de travail. L'application de la *group elicitation method* dans un projet de conception participative a permis, par exemple, de mettre en évidence les problèmes que susciterait l'introduction d'un nouveau dispositif de communication air-sol pour des pilotes (Boy, 1996a).

Enfin, la technique d'élicitation directe (ou associative) est particulièrement recommandée par rapport à d'autres techniques d'élicitation (ex. première hiérarchisation d'un ensemble de produits ou de marques, puis élicitation des critères relatifs à cette hiérarchisation), en raison de sa rapidité d'administration et de sa plus grande congruence avec la façon naturelle des sujets d'évaluer un produit/service (Breivik & Supphellen, 2003).

L'approche systémique de la psychologie appliquée à l'évaluation de la qualité perçue

Nosulenko et Samoylenko, des chercheurs russes contemporains, développent une méthode d'évaluation de la qualité perçue (Nosulenko & Samoylenko, 1997, 1999a) sur la base des travaux des psychologues de l'école russe présentés précédemment. Cette méthode repose sur une prise en compte de l'activité de l'individu dans son contexte socioculturel et de

la médiation que l'environnement ou les artefacts⁴⁸ lui apportent pour former ses représentations (Nosulenko & Samoylenko, 2001).

Au départ, cette méthode est conçue pour étudier les processus psychiques⁴⁹ des sujets (ex. reconnaissance et caractérisation qualitative des sons de divers instruments de musique par des musiciens), puis elle est beaucoup appliquée dans le domaine de l'évaluation subjective/sensorielle des produits/services. Elle a fait ses preuves dans diverses industries, comme l'automobile (bruits de moteur, bruits de claquements de portes) et celle des NTIC⁵⁰ (sites Internet, produits *groupware*⁵¹, environnements augmentés). Elle peut même être employée pour une évaluation comparative de produits concurrents (Nosulenko & Samoylenko, 1998, 2000a, 2000b).

Elle découle de l'Approche Systémique Cognition et Communication qui « *se caractérise par le fait que les processus perceptifs et cognitifs sont analysés conjointement aux processus communicationnels* » (Nosulenko & Samoylenko, 1997). Elle intègre ainsi, dans un même paradigme expérimental, à la fois une méthode d'analyse des communications et une méthode psychophysique (étude des correspondances entre images perceptives et objets, dont diverses caractéristiques sont mesurées objectivement).

L'approche se base, tout d'abord, sur le recueil de verbalisations libres en utilisant l'opération de comparaison, puis sur leur analyse systémique pour dégager des groupes sémantiques correspondant à la perception des sujets. De même que les précédentes techniques, cette méthode cherche à éviter de contraindre les sujets dans le choix de dimensions d'évaluation « pré-choisies ».

- Le recueil des verbalisations libres via l'opération de comparaison :

La technique de recueil des verbalisations libres est analogue à celle des protocoles verbaux dans le sens où il s'agit d'utiliser les informations qui se trouvent dans les verbalisations produites par le sujet alors qu'il perçoit un phénomène ou un objet qu'il utilise. Elle s'en distingue cependant car elle implique que le sujet décrive les caractéristiques des objets perceptifs (Nosulenko & Samoylenko, 1997, 1999a) et non les informations qu'il traite

⁴⁸ Les nouvelles technologies, en particulier.

⁴⁹ Equivalent de processus mentaux, allant des processus sensori-moteurs élémentaires aux processus intellectuels du niveau le plus élevé (Savoyant, 1979).

⁵⁰ Abréviation des nouvelles technologies de l'information et de la communication.

⁵¹ Ou aussi produits du champ du CSCW (Computer Supported Cooperative Work).

au fur et à mesure de l'accomplissement de sa tâche en vue de reconstituer les processus cognitifs liés à cette dernière.

Pour cela, le principe est que le sujet verbalise les stimuli présentés généralement par paires. Deux tâches de comparaison peuvent être proposées au sujet : (a) donner un jugement numérique de la dissemblance entre objets et/ou choisir l'objet préféré, (b) verbaliser la différence et/ou similitude entre les objets. L'opération d'évaluation quantitative permet de mieux se représenter la comparaison entre objets et ensuite de la verbaliser plus facilement (Nosulenko & Samoylenko, 1997). Cependant, la procédure expérimentale comprend les deux tâches réalisées consécutivement seulement dans le cas de stimuli perceptibles rapidement (ex. bruits de moteur de véhicules - Nosulenko *et al.*, 1998, 2000; Nosulenko *et al.*, 1996 - ou timbres musicaux - Samoylenko *et al.*, 1996). Dans le cas de produits/services, cette procédure n'est pas possible : il faut que le sujet utilise un produit/service, puis un second, avant de pouvoir procéder à une comparaison. Or, la nécessité d'une opération de comparaison n'exclut pas la comparaison différée à condition de rapprocher au maximum dans l'espace et dans le temps les produits/services testés, afin que la référence utilisée par le sujet soit explicite dans la situation de test (Nosulenko & Samoylenko, 1999b, 2001). Ainsi, dans le cas de produits/services, seule la seconde tâche de comparaison est donnée à réaliser ainsi que celle du choix « du préféré » (Nosulenko & Samoylenko, 1998, 1999b, 2000a, 2000b).

Concrètement, le sujet est invité à réaliser la même tâche d'utilisation d'un premier produit/service puis d'un second, tout en verbalisant. La consigne est du type : « Pendant que vous explorez la page d'accueil du site Internet, pouvez-vous signaler tout haut, votre impression générale du site, vos remarques, vos appréciations ». A la fin, il est invité à verbaliser les différences et similitudes entre produits/services et à désigner celui qu'il préfère. Des relances neutres (« continuez à dire tout haut vos impressions ») sont utilisées pour ne pas inciter le sujet à des interprétations qui pourraient dénaturer la tâche principale d'utilisation du produit/service.

On retrouve ici la notion de similitude/différence de la technique des triades. En effet, Nosulenko et Samoylenko utilisent, entre autres, les travaux de Lomov et Rubinstein pour montrer que l'opération de comparaison est une opération cognitive majeure qui « *intègre les processus sensoriels et intellectuels* » et « *constitue une forme supérieure de la perception et un élément de base de la pensée* ». « *L'opération de comparaison représente une condition indispensable à tout processus subjectif de mesure, de catégorisation et d'identification des*

objets et évènements » (Nosulenko & Samoylenko, 1997, p. 230-231). La situation de comparaison et donc le protocole expérimental intégrant les opérations de « perception – comparaison – évaluation – verbalisation » représentent ainsi une condition nécessaire pour obtenir les verbalisations pertinentes de l'activité perceptive et cognitive du sujet (Nosulenko & Samoylenko, 1999b).

- Analyse systémique des verbalisations :

Dans un second temps, les verbalisations retranscrites et les enregistrements vidéo synchrones sont analysés. Les verbalisations sont mises en correspondance avec les actions et les opérations du sujet (Lahlou *et al.*, 2002; Nosulenko & Samoylenko, 1999a) et découpées en unités verbales significatives. Une analyse systémique des unités verbales est ensuite réalisée. Elle porte sur trois dimensions : elle « *consiste à définir le sens logique des unités verbales (comment s'effectue la comparaison des objets), leur relation aux objets comparés (quels sont les aspects des objets ayant servi à comparer ceux-ci) et à analyser les aspects sémantiques (quelles significations verbales sont attribuées à chacun des objets comparés)* » (Lahlou *et al.*, 2002; Nosulenko *et al.*, 2000). Les unités verbales sont aussi distinguées selon qu'elles caractérisent l'activité d'usage ou la perception du produit (Nosulenko & Samoylenko, 1997, 2001). Les groupes sémantiques obtenus concernent donc et les propriétés des objets et les caractéristiques opérationnelles de l'activité avec ces objets. Des descripteurs relatifs aux objets et aux opérations sont générés, contrairement aux précédentes techniques.

Un calcul stabilisant la quantité d'unités verbales négatives (ex. « moche ») et positives (ex. beau) sur chaque descripteur permet aussi d'obtenir un portrait verbal. Il s'agit d'une cartographie multidimensionnelle de l'appréciation subjective, d'une part, du produit/service ou d'éléments le constituant (ex. les icônes et menus, l'image, le son ou le Word partagé d'un produit de visioconférence) et, d'autre part, des différentes opérations possibles (ex. manipuler un élément du système, accéder à un service, etc.); tout cela pour une tâche donnée (Nosulenko & Samoylenko, 1998, 2000a, 2000b).

- Les aspects psychophysiques :

La méthode prévoit aussi d'établir un modèle physique des objets (Nosulenko & Samoylenko, 1999a, 2001), c'est-à-dire un ensemble de mesures objectives de leurs principaux paramètres. Il permet d'établir des correspondances entre les données subjectives (données de préférence et portraits verbaux) et les paramètres « objectifs » des objets perçus (Nosulenko & Samoylenko, 2000a). Ces indications permettent d'identifier les paramètres

associés à des jugements défavorables et peuvent déboucher sur des recommandations de conception (Lahlou *et al.*, 2002; Nosulenko & Samoylenko, 1999b, 2001).

Jusque là, nous avons décrit des méthodes d'évaluation qui prévoient la construction de dimensions. Ce n'est pas toujours le cas, des dimensions « prêtes à l'emploi » peuvent être utilisées.

2.6.3. L'évaluation à partir de dimensions déjà construites

Les précédents descripteurs obtenus peuvent être réemployés pour des évaluations ultérieures plus rapides (sous la forme de différentiel sémantique; Nosulenko *et al.*, 1998; Nosulenko & Samoylenko, 1999a, 2001). Il existe ainsi des listes de descripteurs « prêts à l'emploi », dans le domaine des produits informatiques, des NTIC (cf. § 2.7.1.), comme dans celui des produits alimentaires (cafés, chocolats, vins, lait, bières, etc.; Barthélémy, 1998).

Il s'agit néanmoins d'être prudent et d'évaluer des stimuli ou produits/services rigoureusement identiques (Nosulenko *et al.*, 1998; Nosulenko & Samoylenko, 2000a). Dans le domaine des produits alimentaires, il est même recommandé que les sujets élaborent eux-mêmes les descripteurs qu'ils utilisent pour l'évaluation (Urdapilleta, 2001b).

Toutefois, diverses sortes d'échelles peuvent être utilisées pour établir la mesure d'un produit sur une dimension d'évaluation donnée. Dans le domaine de l'alimentaire comme dans l'automobile, diverses recherches sont menées à partir des connaissances constituées dans le champ de la catégorisation et de la logique floue. D'une part, il s'agit de trouver des échelles d'évaluation qui permettent de mieux discriminer les produits (les recherches s'orientent vers les échelles hiérarchiques à items verbaux ; Urdapilleta, 2001c), d'autre part, il s'agit de prendre en compte l'imprécision relative à la représentation mentale de l'échelle que le sujet se forge. La distance « psychologique » entre deux items n'est pas nécessairement constante et les évaluateurs ont recours aux techniques de jugement par intervalle (Les_Cahiers_du_Club_CRIN, 1997; Urdapilleta & Bouard, 2001).

Enfin, nous examinons une dernière forme d'évaluation basée sur des dimensions généralement déjà construites. Il s'agit de questionnaires d'évaluation qui sont souvent « prêts à l'emploi », même si un questionnaire peut aussi être construit, comme nous le verrons dans ce qui suit.

2.7. L'ÉVALUATION PAR QUESTIONNAIRE

Les évaluations par questionnaire sont proches de l'évaluation sensorielle/subjective, puisqu'elles peuvent traiter de dimensions hédoniques et émotionnelles. Néanmoins, elles peuvent aussi interroger des aspects « plus objectifs » (ou du moins, plus faciles à évaluer quantitativement), relatifs à l'utilité et à l'utilisabilité du produit/système.

Par ailleurs, les questionnaires (ou aussi check-lists) auxquels nous faisons référence ne mettent pas à contribution ni les techniques de jugement par intervalle, ni les échelles hiérarchiques à items verbaux. Ils sont généralement construits à partir (a) de questions ouvertes (réponse libre), (b) de questions fermées à choix multiples, (c) de questions scalaires, c'est-à-dire sous forme d'échelles de Likert en 5 ou 7 points (à items verbaux ou numériques, permettant une évaluation bipolaire d'un extrême à un autre), (d) d'un différentiel sémantique (regroupement de plusieurs questions scalaires) ou (e) de questions de rang demandant au sujet de hiérarchiser une liste de produits ou de fonctions selon leur degré d'utilité, de préférence, etc. (Baccino *et al.*, 2005; Dix *et al.*, 1993).

2.7.1. Les questionnaires « prêts à l'emploi »

2.7.1.1. Les questionnaires pour évaluer l'utilisabilité

Un premier ensemble de questionnaires peut être distingué. Il s'agit de questionnaires comme celui de la très connue check-list de Ravden et Johnson (1989) ou comme PUTQ⁵² (Lin *et al.*, 1997). Ils sont destinés à être utilisés par des experts en ergonomie ou tout professionnel ayant développé quelques compétences en matière d'évaluation d'interfaces⁵³. Chacune des questions sert à vérifier l'adéquation des divers éléments de l'interface aux règles ergonomiques générales. Elles sont regroupées en chapitres qui correspondent sensiblement aux heuristiques et aussi aux critères ergonomiques de Bastien et Scapin. Ces questionnaires peuvent ainsi être utilisés à la place des heuristiques dans une inspection experte (Nielsen & Molich, 1990).

⁵² Abréviation de Purdue Usability Testing Questionnaire.

⁵³ Nous pensons, en effet, que PUTQ requière un minimum de connaissances en ergonomie, même si Lin *et al.* (1997) soutiennent que leur questionnaire est conçu pour être utilisé par des utilisateurs.

2.7.1.2. Les questionnaires pour évaluer la satisfaction et/ou l'utilisabilité subjective

Les précédents questionnaires sont orientés vers l'évaluation de l'utilisabilité de l'interface (par un expert) tandis que d'autres cherchent à recueillir le point de vue de l'utilisateur. Leurs questions sont de la forme « J'apprécie de travailler avec ce logiciel – D'accord – Sans opinion - Pas d'accord », « Travailler avec ce logiciel est stimulant sur un plan intellectuel - », « L'organisation de l'information vous semble-t-elle claire ou confuse » (avec une évaluation possible en 7 points). Il s'agit ainsi de l'évaluation subjective de l'utilisabilité de l'interface et de la satisfaction. Aussi qualifiés de questionnaires d'attitudes, ils sont, par exemple : CUSI⁵⁴ (Kirakowski & Corbett, 1988), SUMI⁵⁵ (Kirakowski, 1996), ASQ⁵⁶ (Lewis, 1991), CSUQ⁵⁷ (Lewis, 1995), QUIS⁵⁸ (Chin *et al.*, 1988; Shneiderman, 1998), SUS⁵⁹ (Brooke, 1996).

Ces derniers peuvent soutenir des évaluations à différentes étapes d'un projet de conception ou d'acquisition d'un nouveau logiciel. Par exemple, le questionnaire SUMI, développé dans le cadre de la méthodologie de mesure de performance MUSiC⁶⁰ (Macleod, 1996; Macleod *et al.*, 1997), a été appliqué avec succès dans l'atteinte des trois objectifs suivants : (a) évaluation des logiciels utilisés à différents postes de travail pour identifier ceux qui doivent être remplacés, (b) aide au choix d'un logiciel à acquérir (le questionnaire a permis de rendre les acheteurs attentifs à ce qui doit être pensé avant un tel achat), (c) évaluation comparative d'une nouvelle et ancienne version d'un logiciel, pour savoir si elle peut être commercialisée ou nécessite des améliorations supplémentaires.

Dans le cadre d'un projet de conception, les questionnaires sont administrés aux utilisateurs « réels » après une simple exploration d'un prototype ou du logiciel opérationnel, voire après une séance de test utilisateur. Les questionnaires permettent de recueillir des données sur la satisfaction des utilisateurs, difficiles à obtenir autrement puisque les mesures de performances ou observations comportementales ne fournissent que des données concernant l'efficacité et l'efficacé.

⁵⁴ Abréviation de Computer User Satisfaction Inventory.

⁵⁵ Abréviation de Software Usability Measurement Inventory.

⁵⁶ Abréviation de After-Scenario Questionnaire.

⁵⁷ Abréviation de Computer System Usability Questionnaire.

⁵⁸ Abréviation de Questionnaire for User Interface Satisfaction.

⁵⁹ Abréviation de System Usability Scale.

⁶⁰ Abréviation de Measuring Usability of Systems in Context.

2.7.1.3. Prise en compte de la satisfaction et des dimensions hédoniques et émotionnelles en ergonomie

Il est cependant dénoncé le peu de techniques d'évaluation incluant des mesures de la satisfaction des utilisateurs et le fait que ces mesures se limitent souvent à demander aux utilisateurs ce qu'ils pensent du produit, après avoir relevé un ensemble de mesures de performances (Lindgaard & Dubek, 2002)⁶¹. Il semble que la satisfaction soit davantage un sous-produit de l'utilisabilité (« *usability reductionism* » ; Hassenzahl *et al.*, 2001a), même si l'efficacité, l'efficacités et la satisfaction sont distinguées dans l'utilisabilité (ISO, 1998). Les ergonomes auraient tendance à considérer qu'un produit utilisable est nécessairement satisfaisant et à ne pas distinguer accomplissement et plaisir à faire un travail. De plus, certains trouvent frivole et donc inutile la prise en compte de dimensions hédoniques et émotionnelles dans la conception de produits/systèmes (Hollnagel, 1999).

Pourtant, les émotions jouent un rôle important dans la résolution de problème et dans la prise de décisions (Kim & Moon, 1998) et le plaisir (*joy of use*) est instauré comme critère d'acceptabilité des produits/systèmes (Dejean & Naël, 2004; Jordan *et al.*, 1996b). Ce dernier est même célébré comme *la* dimension des produits du futur (Glass, 1997). Il est aussi démontré qu'en plus de l'utilité perçue, la joie d'utilisation a un effet non négligeable sur la satisfaction et l'intention d'utiliser des technologies informatiques (Igbaria *et al.*, 1994). De même, l'attractivité (*appealingness*) d'un produit est autant influencée par la perception de sa qualité ergonomique que par celle de sa qualité hédonique (Hassenzahl, 2001; Hassenzahl *et al.*, 2000b). Des corrélations ont même pu être établies entre des impressions esthétiques complexes (ex. la confiance, la bizarrerie, l'élégance) et des éléments visuels d'interfaces (Kim & Moon, 1998; Park *et al.*, 2004).

Plus globalement, des dimensions relatives au développement personnel (ex. besoin d'être stimulé par la nouveauté ; Hassenzahl, *et al.*, 2000a; Sutcliff, 2000) et à l'aspect social (ex. le besoin de dire qui on est à travers les objets ; Leventhal *et al.*, 1996) interviennent dans l'acquisition et l'utilisation d'un produit (Hassenzahl *et al.*, 2001a). En effet, il ne faut pas oublier que l'expérience humaine recouvre des dimensions liées au corps et à l'esprit (cognition, émotion, volonté) (Rijken & Mulder, 1996). Cela rejoint aussi la psychologie de tradition russe où l'attitude et les motifs d'un sujet interviennent dans son interaction avec le monde et l'orientent.

Pour conclure, la tendance en ergonomie est globalement orientée vers la prise en compte des dimensions hédoniques (Jordan *et al.*, 1996b), même s'il reste du chemin à parcourir (ex. Lin *et al.*, 1997). Auparavant l'ergonomie cherchait seulement à quantifier le niveau de satisfaction des utilisateurs et à réduire les effets négatifs (fatigue, inconfort, charge mentale, etc.) alors que maintenant elle veut contribuer à augmenter les effets positifs, au niveau émotionnel en particulier.

2.7.1.4. Variété des dimensions d'évaluation dans le domaine du Web

Cette évolution est sans doute à rapprocher de l'émergence des sites Internet, puisque « *les activités ne sont [alors] plus exclusivement liées au travail, mais aussi au loisir, à la culture et à l'acquisition de biens* » (Leulier *et al.*, 1999, p. 49). Les dimensions d'évaluation de ce domaine sont plus variées qu'en ergonomie « traditionnelle » et définitivement orientées vers les aspects hédoniques et émotionnels (Hassenzahl & Trautmann, 2001; Kim & Moon, 1998; Lindgaard & Dudek, 2002; Park *et al.*, 2004).

En termes de questionnaires dans le domaine du Web, on peut citer WAMMI⁶² (Kirakowski & Claridge, 2001; Kirakowski *et al.*, 1998), adapté du questionnaire SUMI et mesurant les mêmes dimensions. Dans le domaine du Web marchand, le questionnaire SERVQUAL⁶³ (Parasuraman *et al.*, 1985, 1991, 1993), issu du marketing, est très fréquemment utilisé pour évaluer la qualité perçue du service (la perception de la fiabilité du service rendu, de l'attention personnalisée du service, etc. ; Jarvenpaa & Todd, 1997; Kim & Lee, 2002; Liu & Arnett, 2000). La plupart du temps, les évaluateurs ont incorporé le questionnaire à des méthodes ergonomiques. Il est administré non pas via Internet ou par courrier postal mais après que les sujets aient réalisé plusieurs tâches d'utilisation d'un site Internet marchand. Il est, en effet, préconisé que les questionnaires soient précédés d'une phase d'utilisation du produit/système (Nielsen, 1993).

Enfin, de nombreux questionnaires sont faits « sur mesure » et convoquent des dimensions d'évaluation spécifiques aux objectifs de l'étude, dans le domaine du Web marchand (Jarvenpaa & Todd, 1997; Jarvenpaa *et al.*, 2000; Konradt *et al.*, 2003; Lee *et al.*, 2000; Liu & Arnett, 2000; Teo, 2002) et aussi dans tout autre domaine. Dans ce qui suit, nous

⁶¹ Excepté le questionnaire SUMI qui comprend six échelles : *Global, Affect, Efficiency, Helpfulness, Control* et *Learnability*. SUMI permet donc de mesurer la réaction émotionnelle de l'utilisateur et sa perception quant au niveau d'efficacité, de serviabilité, de contrôle et de facilité d'apprentissage que lui offre le logiciel. L'échelle *Global* correspond au score général qui reflète l'utilisabilité subjective globale.

⁶² Abréviation de Web Analysis and Measurement Inventory.

⁶³ Réduction de Service Quality.

abordons les études exploratoires dont le questionnaire et les dimensions d'évaluation sont créés spécifiquement pour l'étude. Nous décrivons aussi leur processus de construction.

2.7.2. Les études exploratoires et la construction d'un questionnaire

2.7.2.1. Les études exploratoires

Les études exploratoires par questionnaire sont quasiment de type sociologique et cherchent à recueillir des informations sur les utilisateurs pour améliorer les fonctionnalités et aspects d'un produit. Elles peuvent être réalisées en phase amont de la conception (Baccino *et al.*, 2005; Maguire & Bevan, 2002; Rubin, 1994) ou en phase de « feed-back », une fois le produit/système sur le marché, de même que les entretiens individuels. Toutefois, les entretiens individuels offrent une plus grande liberté pour s'adapter à la compréhension du sujet et à l'évolution des questions de recherche (Blanchet & Gotman, 1992).

Comme nous le verrons dans cette thèse, les études par questionnaire peuvent aussi avoir pour objectif une activité de veille type « benchmark », concernant les produits du grand public notamment où le marketing et l'ergonomie travaillent davantage main dans la main (Dejean & Naël, 2004; Leulier & Donzier, 2002). Les résultats des études constituent alors des indicateurs quant à l'usage et à l'appréciation qualitative d'une famille de produits et de leurs concurrents. Ces indicateurs peuvent guider les chefs de projet dans leurs diverses prises de décision.

Les études « GVU's WWW User Surveys »⁶⁴, menées semestriellement depuis 1994 par questionnaire en ligne, en constituent un bon exemple (Kehoe *et al.*, 1997; Kehoe *et al.*, 1998a; Kehoe *et al.*, 1998b; Pitkow & Kehoe, 1996). Ces études ont permis de mettre au jour un grand nombre de données sur les internautes américains et européens. Au-delà des données démographiques recueillies, ces études révèlent les attitudes et/ou usages des internautes quant aux cookies⁶⁵, aux transactions par carte bancaire, aux spams⁶⁶, aux lois informatiques en vigueur, etc. Elles permettent aussi de savoir quel navigateur est le plus utilisé ou

⁶⁴ Etudes sur les internautes et cyber-acheteurs (américains et européens) menées par le Graphics Visualization and Usability Center, Georgia Institute of Technology en partenariat avec le W3C (World Wide Web Consortium) et l'INRIA.

⁶⁵ Petit fichier enregistré sur le micro-ordinateur de l'internaute lors de sa première visite sur un site Internet, pour enregistrer des données sur son parcours et éventuellement personnaliser le site (Bonjour Monsieur... ; Fouchard, 2001).

⁶⁶ Courriels publicitaires envoyés souvent sans l'accord de l'internaute dans sa boîte de messagerie.

d'apprendre les principales raisons des internautes pour accepter ou au contraire rejeter l'achat par Internet.

2.7.2.2. La construction d'un questionnaire

Pour construire ces questionnaires d'étude, tout un processus est généralement suivi. La littérature sur l'enquête en sociologie et en psychologie sociale nous renseigne sur ce point. Une phase qualitative amont (ou pré-enquête) est généralement menée pour définir les objectifs de l'étude et recueillir le matériel nécessaire à la formulation du questionnaire. Il s'agit de se documenter à partir des travaux déjà menés sur le thème, puis de recueillir les connaissances des « spécialistes » du domaine étudié ou les points de vue de répondants potentiels lors de séances de *brainstorming*, d'entretiens en groupe ou d'entretiens approfondis (Blanchet & Gotman, 1992; Ghiglione & Matalon, 1998; Javeau, 1982; Morgan, 1997; Mucchielli, 1968, 1996).

Une fois définies les hypothèses et dimensions d'étude, il s'agit de déterminer le mode d'administration du questionnaire (par Internet, par courrier, par entretien dirigé, voire par téléphone), la taille de l'échantillon à extraire de la population mère et enfin les techniques d'échantillonnage (Javeau, 1982). Une ébauche de questionnaire est ensuite construite dans un processus itératif de rédaction de projets de questionnaire et de critique par des personnes compétentes, avant d'aboutir à un pré-test où la mouture finale est testée auprès de répondants potentiels. Après remaniement si nécessaire, le questionnaire est alors prêt à l'emploi.

Des règles de construction de questionnaire ont été élaborées dans de nombreux manuels (Ghiglione & Matalon, 1998; Javeau, 1982; Mucchielli, 1968) et concernent (a) le choix du type de questions (questions ouvertes ou fermées, questions de type vrai/faux ou avec réponses intermédiaires, etc.), (b) la formulation des questions (un questionnement impliquant directement le répondant ou le contraire, le choix du vocabulaire, l'attention aux références symboliques comme les noms de personnalité, etc.), (c) la structure et l'ordre de succession des questions (la progression dans les questions pour obtenir un effet « entonnoir », la dispersion ou au contraire les redondances de questions pour le contrôle des réponses, etc.) et enfin (d) la longueur du questionnaire.

Le choix du type de questions (ouvertes/fermées) doit correspondre notamment au degré de précision informationnelle souhaité, mais aussi aux moyens de traitement disponibles. La plupart des autres règles correspondent au souci d'éviter les déformations involontaires, à savoir la réaction de prestige ou « de façade », la contraction défensive à la question

personnalisée, l'effet de halo ou de contamination des questions les unes par les autres, etc. (Javeau, 1982; Mucchielli, 1968). Il s'agit aussi d'éviter tout écart sémantique entre le chercheur (ou expert) et le répondant (Kitzinger, 1994; Morgan, 1997; Nosulenko *et al.*, 1998; O'Brien, 1993). La difficulté dans la formulation des questions ou échelles de jugement réside, en effet, dans le fait de ne pas engendrer de contresens (Fuller *et al.*, 1993; Urdapilleta, 2001b).

Une fois l'ébauche finale du questionnaire obtenue, il s'agit enfin de la tester pour vérifier si elle ne provoque ni incompréhension ni réaction de déformation ni désintérêt. A ce stade, un pré-codage ou plan de dépouillement (prévoir toutes les opérations à réaliser sur les réponses une fois recueillies) peut être prévu et permet de s'apercevoir d'éventuels oublis ou de redondances (Ghiglione & Matalon, 1998). Par ailleurs, après le pré-test, la validité du questionnaire doit normalement être mise à l'épreuve en procédant à plusieurs itérations de test statistique. Il s'agit de procéder à une analyse statistique des données, après avoir recueilli les réponses de participants (Brooke, 1996; Kirakowski, 1996). Des questions peuvent notamment être supprimées quand elles contribuent à diminuer le degré de fiabilité du questionnaire (Kim & Lee, 2002).

Dans cette thèse, nous décrivons un processus de construction d'un questionnaire légèrement différent de ce processus « traditionnel » et dans le champ de l'ergonomie.

3. LES QUESTIONS DE RECHERCHE

Nous y abordons aussi la question relative à l'origine des descripteurs ou dimensions d'évaluation. Les débats dans le champ de l'évaluation ergonomique concernent la comparaison des dimensions élaborées par des experts et celles élaborées par des sujets naïfs. Les questions posées sont les suivantes : des dimensions utilisateurs sont-elles plus valides car issues de données obtenues auprès de sujets directement concernés par l'activité étudiée ? Ou bien, les utilisateurs ne sont-ils pas à même de les produire, s'attachant à des détails non pertinents, en raison de la situation artificielle d'élaboration des dimensions ? D'après Piaget, il y a possibilité de réussite sans la conceptualisation. De même, les utilisateurs peuvent choisir, préférer, utiliser des produits/systèmes sans nécessairement avoir fait la démarche de prise de conscience et d'analyse des raisons de ces choix. Les experts sont-ils donc plus à même de les produire grâce à leur recul, expérience et connaissances des utilisateurs ?

Nous évoquons également la question relative au nombre de sujets nécessaires. Cette dernière met en balance la validité et la productivité des données ou résultats d'évaluation. Quel nombre de sujets est nécessaire afin de ne pas sacrifier des aspects de la réalité de l'activité (spécificités des tâches d'utilisation, différences interindividuelles en termes de détection d'un problème mais aussi, plus généralement, d'usages spécifiques) ? D'autre part, quel est le minimum possible de sujets pour obtenir des résultats et valides et les plus « rentables » (production d'informations non redondantes pour ne pas réaliser de dépenses inutiles) ? Autrement dit, quelle est la productivité d'une méthode ?

Enfin, nous traitons des questions relatives à la comparaison de deux techniques présentées ci-dessus : l'entretien individuel et le *focus group*. Des questions sont posées concernant la nature et l'exploitabilité de leur production ainsi que concernant leur productivité dans le champ de l'évaluation ergonomique (et d'autres disciplines également).

Ces trois précédents éléments constituent des débats et questions posés dans le champ de l'évaluation ergonomique. Nous détaillons les questions de recherche opérationnelles dans le chapitre suivant. De plus, nous y étudions la question de l'utilisation des résultats, à laquelle nous avons seulement fait allusion jusque là.

CHAPITRE 2

LES DEUX VERSANTS DE L'ÉVALUATION ERGONOMIQUE : PRODUCTION DE RESULTATS ET LEUR UTILISATION

1. LA QUESTION DE LA PERTINENCE D'UNE EVALUATION

Dans le précédent chapitre, nous avons exposé les principales préoccupations dans le domaine de l'évaluation des produits/systèmes et l'état de l'art des méthodes et des techniques. Nous aborderons ici la question de la pertinence d'une évaluation. Une bonne évaluation doit répondre à deux types de critères qui ne sont pas nécessairement convergents : ceux relatifs à la production de résultats scientifiquement valides, d'une part, et ceux relatifs à l'utilisation des résultats par les professionnels destinataires, d'autre part. Ceci nécessite de prendre en compte les deux « versants » de l'évaluation ou les deux activités déployées autour de l'évaluation ; production de résultats puis leur utilisation.

La première partie de cette thèse s'inscrit dans le premier « versant ». Ce travail a pour origine une demande sociale au sens d'une demande issue du milieu industriel (et non d'une demande issue d'organismes publics ; Spérandio, 1985). Les professionnels du service d'évaluation de France Télécom, pris dans une conjoncture économique de plus en plus compétitive, cherchent à mieux estimer et à augmenter l'efficacité des méthodes d'évaluation qu'ils emploient. Ils nous ont demandé de comparer la pertinence de deux techniques utilisées dans la phase qualitative amont à la construction d'une évaluation par questionnaire : l'entretien individuel et le *focus group*. Une étude comparative des deux situations d'interaction verbale mobilisant d'autres techniques d'élicitation a été menée. Ces techniques et leurs auteurs sont précisés au chapitre précédent. Il s'agit de la technique d'élicitation associative, de l'entretien d'explicitation et de la méthode des verbalisations libres utilisée dans l'approche systémique de la psychologie appliquée à la qualité perçue. Cette étude porte sur la perception des utilisateurs de sites Internet marchands.

Dans la seconde partie de cette thèse, un déplacement de point de vue est opéré et la perspective est élargie à l'ensemble des acteurs intervenant dans le processus d'évaluation et de conception. En effet, s'il y a besoin d'obtenir des résultats ergonomiques valides dans le sens de refléter la réalité des utilisateurs des produits/systèmes, qu'en est-il de l'utilisation de ces résultats ? Un « bon » résultat du point de vue d'un ergonome est-il aussi nécessairement « bon » du point de vue des professionnels qui vont les utiliser et transformer les produits/systèmes ? Qu'est-ce qu'un « bon » résultat ergonomique : celui qui relate une connaissance très précise de l'utilisation des produits/systèmes ou celui qui permet une prise de décisions pour la transformation et l'amélioration du produit/système ? D'une logique centrée exclusivement sur la production de dimensions utilisateurs et de résultats

ergonomiques, nous passons à une logique de l'utilisation des résultats par les professionnels de la transformation⁶⁷ des produits/systèmes.

Cette seconde partie de thèse consiste en un travail d'analyse des besoins en connaissances ou « besoins en évaluation » de professionnels du Web marchand (veille-marketing, marketing, conception) et en une activité de conception d'un artefact instrumentalisable par ces derniers. La nécessité d'un outil d'aide à l'élaboration d'évaluations s'est imposée car certains de ces professionnels sont, non seulement utilisateurs de résultats d'évaluation, mais aussi producteurs de ces résultats et concepteurs de leurs propres évaluations, de même que les professionnels du service d'évaluation déjà évoqués.

Dans ce chapitre, nous mobilisons tout d'abord les connaissances relatives aux questions soulevées par les professionnels du service d'évaluation concernant la production de dimensions subjectives et de contenu verbal pour la construction d'un questionnaire. Nous abordons ensuite le renversement de points de vue sur l'évaluation et présentons l'approche instrumentale (Rabardel, 1995) qui soutient ce mouvement et peut être une base pour la conception d'un artefact utilisable, c'est-à-dire instrumentalisable par des sujets. Enfin, nous détaillons les terrains et notre démarche de recherche.

2. LES CONNAISSANCES SUR L'ELABORATION DE DONNEES SUBJECTIVES ET VERBALES POUR LA CONSTRUCTION D'UN QUESTIONNAIRE

Les questions de recherche de la première partie de nos travaux sont étroitement liées à l'activité des auteurs de la demande émanant du terrain. Nous avons réalisé l'étude comparative systématique de l'entretien individuel et du *focus group*⁶⁸ pour la construction d'un questionnaire sur la base des interrogations des professionnels du service d'évaluation, préoccupés par l'efficacité de leurs méthodes.

⁶⁷ Transformation renvoie à toute l'activité de spécification, de réalisation, puis de transformation successive d'un produit/système (d'une version à une autre) ; ceci aussi bien sur le plan technique que marketing.

⁶⁸ Nous utilisons ici le terme américain *focus group* là où les francophones emploient plus généralement le terme « groupe d'animation », le terme *focus group* étant réservé à une utilisation de ce type d'entretien dans le domaine du marketing. Dans notre démarche, l'entretien en groupe est utilisé comme technique d'élicitation d'idées dans un cadre de qualité perçue, qu'il s'agisse de dimensions ergonomiques, hédoniques ou relatives à la qualité de service (comprenant des aspects marketing).

L'ensemble de leurs interrogations porte sur :

- 1. la pertinence de dimensions d'évaluation produites à partir de données recueillies auprès d'utilisateurs de produits/services (dimensions utilisateurs) par rapport à la pertinence de dimensions d'évaluation produites par des experts à partir de leurs connaissances des utilisateurs/clients (dimensions experts),
- 2. la pertinence de l'entretien individuel versus du *focus group* concernant la nature et l'exploitabilité des contenus verbaux produits (permettant l'élaboration de dimensions d'évaluation),
- 3. la pertinence de l'entretien individuel versus du *focus group* en termes de productivité par rapport aux coûts financiers et aux ressources humaines nécessaires à leur réalisation,
- 4. l'intérêt de réaliser un entretien supplémentaire en termes de productivité, c'est-à-dire savoir à partir de quel nombre d'entretiens ou de sujets, on observe un effet de saturation ou de stagnation de la productivité,
- 5. la pertinence de la production verbale (nature, exploitabilité) lors d'une situation de verbalisation avec utilisation simultanée du produit/service versus sans utilisation.

Chacune de ces interrogations renvoie à des éléments de connaissances déjà constituées ou à des questions posées dans le champ de l'évaluation ergonomique que nous développons dans les paragraphes 2.1. à 2.4. ci-dessous.

2.1. DIMENSIONS D'ÉVALUATION PRODUITES PAR DES UTILISATEURS VERSUS PAR DES EXPERTS

Comme nous l'avons signalé, les résultats de prédiction de l'utilisabilité d'un produit/système issus de méthodes expertes d'inspection sont souvent comparés aux résultats empiriques obtenus lors de tests utilisateurs. Ces derniers sont considérés plus valides car faisant intervenir des utilisateurs dans un contexte proche de la situation naturelle (généralement dans un laboratoire aménagé à cet effet).

Dans le domaine de l'évaluation des produits alimentaires, ce sont plutôt les données recueillies auprès de sujets naïfs qui sont considérées peu valides par rapport à des données produites par des sujets entraînés à l'évaluation (ex. méthode de profil conventionnel - Barthélémy, 1998 ; méthode basée sur un travail de groupe pour la génération et la définition collectives de descripteurs - Urdapilleta, 2001b). En effet, « demander aux consommateurs de décrire les caractéristiques sensorielles du produit non seulement fournira des données

analytiques de mauvaise qualité, mais conduira aussi à des distorsions dans leurs jugements hédoniques. Cette description sensorielle du produit doit être effectuée par des groupes spécialement entraînés » (Köster, 1998b, p. 200) et « *il peut être beaucoup plus économique de questionner les sujets, mais les réponses qu'ils fournissent certes avec facilité ne correspondent pas toujours à leur comportement réel* » (Köster, 1998a, p. 183). Ainsi, il est nécessaire d'avoir d'abord recours à des sujets entraînés pour accomplir l'évaluation d'une grandeur sensorielle complexe (évaluation multidimensionnelle) des produits. Ces données peuvent ensuite être corrélées à des données sur la consommation de ces produits ou à des mesures hédoniques simples⁶⁹ recueillies auprès de sujets non entraînés.

En matière d'élaboration d'hypothèses ou de dimensions d'évaluation pour une enquête par questionnaire, des sujets naïfs comme des sujets spécialistes peuvent aussi être interrogés (entretiens, *focus groups*, brainstormings, etc.). En transposant les précédents constats à notre problématique, on peut s'interroger, de même, sur l'intérêt de dimensions élaborées par des experts par rapport à des dimensions élaborées à partir de données utilisateurs (et inversement). Les deux techniques avec les utilisateurs apportent-elles des éléments nouveaux et donc utiles pour l'identification de dimensions d'évaluation, par rapport à une liste de dimensions élaborées par des experts à partir de leurs connaissances des utilisateurs/clients, de leurs usages, de leurs difficultés, de leurs attentes et exigences, etc. ?

Après avoir examiné la question relative à la comparaison des données utilisateurs versus experts, nous abordons celles relatives à la comparaison de l'entretien individuel et du *focus group*.

2.2. COMPARAISON DE L'ENTRETIEN INDIVIDUEL ET DU *FOCUS GROUP* POUR LA CONSTRUCTION D'UN QUESTIONNAIRE

Après avoir constaté la rareté des travaux comparant les deux techniques utilisées pour la construction d'un questionnaire, nous exposons de façon détaillée des éléments de connaissances issus de divers domaines d'application concernant la nature et l'exploitabilité de la production verbale des deux techniques ainsi que concernant leur productivité.

⁶⁹ L'évaluation hédonique est destinée à estimer dans quelle mesure le produit est apprécié lors d'une première consommation, sur le long terme, etc. Pour ce faire, seules des questions hédoniques simples doivent être posées et diverses épreuves ont été mises au point (ex. épreuve de classement d'échantillons, de consommation en situation « naturelle simulée », etc.).

2.2.1. Les connaissances actuelles

Les forces et faiblesses de l'entretien individuel et du *focus group* sont relativement bien connues en ergonomie ou en conception de produits/systèmes (Baccino *et al.*, 2005; Caplan, 1990; Kuhn, 2000; Maguire, 2003; O'Donnell *et al.*, 1991; Urban & Hauser, 1993; Van Vianen *et al.*, 1996), d'une part, et en sciences sociales appliquées et/ou en marketing (Crabtree *et al.*, 1993; Gordon & Langmaid, 1988; Morgan, 1997; Stewart & Shamdasani, 1990), d'autre part, même si les avis divergent parfois.

En revanche, en matière d'usage de ces techniques pour la construction d'un questionnaire, peu de connaissances sont disponibles sur leurs avantages et désavantages (si ce n'est en transposant celles évoquées précédemment). L'utilité du *focus group* a été assez largement explorée pour la formulation de questionnaires, en particulier en sciences sociales appliquées (Barrett & Kirk, 2000; Fuller *et al.*, 1993; Morgan, 1997; O'Brien, 1993), mais peu d'études comparent le *focus group* et l'entretien individuel utilisés dans cet objectif.

Ainsi, le *focus group* peut aider la préparation et mise en œuvre de l'enquête (choix de segments de population à étudier, faire connaître l'enquête auprès de la population recherchée, etc.) tout autant que la formulation du questionnaire. Concrètement les *focus groups* peuvent aider (a) à identifier les différents domaines qui doivent être mesurés, (b) à déterminer les hypothèses de travail ou dimensions qui constituent ces divers domaines, (c) à fournir les formulations adéquates pour que les répondants saisissent précisément le sens des questions posées, (d) à comprendre des pratiques ou habitudes qui renseignent directement les modèles comportementaux à étudier. Or, les entretiens individuels comme les entretiens en groupe sont utiles pour fournir ce type de données (Blanchet, 1992; Ghiglione & Matalon, 1998; Javeau, 1982; Mucchielli, 1968, 1996). Ce que nous ne savons pas précisément est quelle est la nature du contenu verbal produit (ex. détails ou au contraire points essentiels), quelle est la difficulté à exploiter les données recueillies et quelle est la productivité (en termes de temps, de coûts, de ressources humaines) de chaque technique.

A notre connaissance, seuls les travaux de Fern (1982) comparent de façon systématique le contenu verbal produit (nombre et qualité des idées) avec l'une et l'autre technique avec une tâche de génération d'idées⁷⁰. Les autres études comparatives concernent la génération de

⁷⁰ Ces idées pouvant servir à la construction d'un questionnaire, même si les conclusions de cette étude portent sur la possibilité de renseigner des concepts assez complexes comme la perception d'un objet, éventuellement symbolique (ex. l'expansion du rôle de la femme dans le milieu militaire).

besoins utilisateurs/clients (Griffin & Hauser, 1993) ou le développement de symboles picturaux (Macbeth *et al.*, 2000) et se rapportent donc à d'autres domaines d'application.

Les travaux de la première partie de cette thèse cherchent à mieux caractériser la production verbale produite par l'une et l'autre technique en vue de la construction d'un questionnaire. Pour cela, nous mobilisons d'abord toutes les connaissances disponibles sur les deux techniques, quels que soient la discipline et l'objectif pour lesquels elles sont utilisées. Par ailleurs, en raison de la plus grande « jeunesse » du *focus group*, les travaux le concernant cherchent souvent à établir son intérêt par rapport à celui de l'entretien individuel ou d'autres techniques d'évaluation qualitative, dont les avantages sont déjà avérés. Aussi dans ce qui suit, les éléments de comparaison se rapportent souvent au *focus group*.

Ils concernent :

- la nature des contenus produits dans l'une et l'autre technique,
- la difficulté d'exploitation des données produites,
- la productivité en termes de temps et de coûts de l'une et de l'autre technique.

2.2.2. Nature des contenus produits

Nous décrivons les caractéristiques connues des entretiens individuels et celles des *focus groups* pour dégager des dimensions de comparaison de la nature des contenus produits par chaque technique. Ces dimensions concernent :

- le niveau de qualité ou de détail des données produites,
- leur spécificité (les données spécifiques à une technique, c'est-à-dire qui n'émergent pas avec l'autre technique),
- et enfin la variété des différents aspects abordés et les points essentiels dégagés par chaque technique.

2.2.2.1. Qualité ou niveau de détail des données

Les caractéristiques de l'entretien individuel

L'entretien individuel est une solution plus adaptée pour interagir avec des personnes (a) très timides, (b) présentant des difficultés à s'exprimer en public et à échanger des informations avec d'autres personnes (ex. les personnes très âgées ; Barrett & Kirk, 2000), (c) ayant des raisons pour ne pas vouloir divulguer certaines informations publiquement ou (d) présentant des différences culturelles/hierarchiques trop fortes pour interagir ensemble (Baccino *et al.*, 2005; Morgan, 1993, 1997). En effet, la situation individuelle offre un plus

grand calme et une plus grande confidentialité. Elle permet de mieux approfondir une expérience ou un parcours individuel. L'interviewé peut aussi s'exprimer plus longuement sur un thème donné. Lors d'un *focus group*, chaque participant aura seulement le temps d'exprimer une certaine part des informations qu'il aurait dévoilées lors d'un entretien en face à face de même durée (Morgan, 1997; Urban & Hauser, 1993). Par ailleurs, l'intervieweur contrôle davantage le questionnement. Par exemple, il peut investiguer en profondeur un besoin d'un utilisateur/consommateur par un questionnement adroit et minutieux (Urban & Hauser, 1993).

Ainsi, le *focus group* fournirait des informations moins approfondies (Morgan, 1997; Van Vianen *et al.*, 1996) et l'entretien individuel davantage de meilleures idées (Fern, 1982)⁷¹. Ce point de vue est néanmoins à nuancer, comme nous le voyons dans ce qui suit

Les caractéristiques du focus group

Le *focus group* présente des caractéristiques complexes qui peuvent avoir des effets positifs, mais aussi parfois négatifs. Sa principale caractéristique est d'être interactif et synergique contrairement aux autres techniques d'évaluation qualitative (ex. Langford & McDonagh, 2003a; Stewart & Shamdasani, 1990). Il présente des différences en termes (a) d'objectifs pour lesquels il est le plus approprié, (b) de conduite d'entretien et (c) de dynamiques dans l'entretien.

Le *focus group* est plus approprié pour approfondir certaines expériences soit très routinières, soit peu réfléchies par avance, pour lesquelles un sujet serait peu loquace pendant une heure d'entretien en face à face. Il est nécessaire, par exemple, d'avoir recours au *focus group* lorsqu'il s'agit de faire discuter des professeurs et étudiants sur leurs usages d'une savonnette (Morgan, 1997) ! D'ailleurs, des *focus groups* peuvent être tenus au domicile des participants, en ergonomie comme en sciences sociales (Krueger & Casey, 2000; Maguire, 2003). Cependant, le *focus group* est aussi utile pour faire connaître et remonter des informations concernant des populations minoritaires, défavorisées ou handicapées (Krueger & Casey, 2000; Maguire, 2003; Morgan & Krueger, 1993) et pour traiter certains thèmes sensibles comme le sida (Kitzinger, 1994) ou l'homosexualité (O'Brien, 1993).

Par ailleurs, l'animateur d'un groupe est partagé entre laisser libre cours aux interactions de groupe (ce qui peut faire émerger des informations inattendues) et recadrer les discussions

⁷¹ La conclusion de cette étude ne tient cependant pas compte des durées des entretiens comparés. Le même nombre de participants est simplement mobilisé pour l'une et l'autre technique.

qui commencent à devenir hors sujet (Morgan, 1997; Nielsen, 1993). Il est pris dans ce dilemme de façon plus prégnante que dans l'entretien individuel.

De plus, les participants d'un *focus group* subissent une certaine pression sociale engendrant des réactions de conformité ou au contraire de provocation (Wight, 1994). Il peut y avoir un effet de leadership où un ou plusieurs participants influencent particulièrement les points de vue des autres, ce qui peut dénaturer les opinions recueillies (Baccino *et al.*, 2005; Langford & McDonagh, 2003a; Mucchielli, 1996). Une trop forte implication de certains participants est donc néfaste, mais l'est également une trop faible implication où les participants ne se sentent pas concernés et à qui il faut extirper des informations qui, de ce fait, sont moins valables (Morgan & Krueger, 1993).

Un inconvénient des *focus groups* est ainsi de ne pas pouvoir prévoir à l'avance quelles seront la dynamique de groupe et la qualité des données recueillies. La technique de conduite de l'animateur peut largement contribuer à diminuer ces effets mais elle ne les supprime pas (Langford & McDonagh, 2003a; Mucchielli, 1996).

En revanche, le groupe forme une entité face à l'animateur (Mucchielli, 1996) et un sentiment de sécurité peut être procuré aux participants du fait de découvrir qu'ils partagent des sentiments et expériences similaires (Krueger & Casey, 2000; Langford & McDonagh, 2003a; Merton *et al.*, 1956; Stewart & Shamdasani, 1990). Entendre les autres les aide à identifier à quel point ce qu'ils entendent correspond à leur situation ou non. Avec les diverses comparaisons et contrastes, les participants deviennent plus explicites (Morgan & Krueger, 1993; Mucchielli, 1996). Les interactions de groupe offrent l'opportunité de dénoncer certains énoncés sur idéalisés (Morgan, 1993), de remettre en question et même de modifier leurs propres opinions de départ, ce qui signifie qu'elles s'affinent et se clarifient au cours des échanges (Kitzinger, 1994).

Le *focus group* est ainsi considéré comme un bon outil pour étudier les motivations complexes, les représentations inconscientes, non rationnelles et sous-jacentes aux jugements, particulièrement dans le domaine des sciences sociales (Kitzinger, 1994; Morgan, 1997; Morgan & Krueger, 1993; Mucchielli, 1996), mais aussi dans le domaine de la conception de produits/systèmes (Baccino *et al.*, 2005). Néanmoins, l'entretien individuel non directif est aussi approprié pour explorer la structure du cadre de référence d'un sujet et l'intensité de ses affects concernant un problème (Ghiglione & Matalon, 1998). Il est reconnu comme un outil pour enquêter sur les comportements complexes en sciences sociales comme en marketing (Blanchet, 1992; Ghiglione & Matalon, 1998; Gordon & Longmaid, 1988).

Il n'est donc pas évident d'établir précisément quelles sont les différences en termes de nature de données obtenues avec l'une ou l'autre technique. Nous proposons de distinguer les deux techniques en termes de niveau de détail des données qu'elles produisent. Cela permettra notamment de vérifier s'il est juste de penser que l'entretien individuel fournit davantage de données approfondies (Fern, 1982; Morgan, 1997).

En dehors du niveau de détail des données produites, nous proposons de chercher à savoir si l'une ou l'autre technique produit des données qui n'émergent pas avec l'autre technique et qui offrirait un contenu spécifique.

2.2.2.2. Données spécifiques à chaque technique

Le *focus group* est réputé pour apporter des données qui n'émergent tout simplement pas dans l'entretien individuel ou dans d'autres techniques en passation individuelle (Agar & MacDonald, 1995; Ede, 2005; Kitzinger, 1994; Krueger & Casey, 2000; Kuhn, 2000; Langford & McDonagh, 2003a; O'Donnell *et al.*, 1991). Par exemple, des agriculteurs dévoilent les raisons qui les feraient venir à des séances de formation lors de *focus groups* alors qu'elles n'ont jamais été percées à jour lors d'un premier processus d'enquête (Krueger & Casey, 2000). De même, des besoins utilisateurs pour la conception d'un système d'automatisation domestique, liés très concrètement au contexte d'utilisation (inventaire du contenu du réfrigérateur, protection automatique contre les pannes de courant, etc.), sont établis et n'auraient pu l'être lors d'entretiens individuels ou d'enquêtes par questionnaire (Kuhn, 2000). Enfin, les *focus groups* peuvent servir à étudier des questions qui portent sur le long terme et seraient trop coûteuses à étudier sur le terrain. De plus, ils peuvent révéler certains aspects des métiers que d'autres techniques ne révèlent pas, notamment car on peut donner l'occasion aux participants de raconter et d'interagir à partir de leur journée de travail (ex. le métier d'administrateur système, au premier abord très technique, est en fait très prenant au niveau relationnel et en termes de documentation ; Ede, 2005).

L'effet synergique des interactions conduit le groupe à produire des idées et récits qui n'auraient pas été énoncés autrement. Il peut s'agir d'idées que l'intervieweur en face à face n'aurait simplement pas songées à explorer (Agar & MacDonald, 1995) ou d'idées qui émergent lors de digressions presque hors sujet (Kitzinger, 1994).

Même la dimension conflictuelle des interactions d'un groupe peut présenter un avantage en ce sens (Kitzinger 1994). Effectivement, si les effets de cohésion sont bénéfiques à un groupe (ex. augmentation du nombre global des interactions), les groupes conflictuels

présentent aussi certains avantages : « *On observe, ainsi, que des groupes potentiellement conflictuels réussissent mieux que des groupes cohésifs dans des tâches d'élaboration complexe. Les divergences d'opinions, les conflits et polémiques protègent le groupe contre les prises de décisions positives conventionnelles et normatives, mais entraînent un biais d'extrémisation des opinions et des décisions [...]* » (Blanchet & Trognon, 1994, p. 88).

Nous proposons d'observer si les interactions de groupe occasionnent des idées qui n'émergent pas en entretien individuel et inversement. Par ailleurs, nous suggérons d'analyser comment sont produites ces idées : particulièrement lors de discussions de confrontation ou non.

L'intérêt de l'utilisation de ces techniques ne réside pas uniquement dans leur capacité à produire des données qui leur sont spécifiques ou des données approfondies. Il est aussi pertinent de vouloir les employer pour dégager une variété de dimensions d'évaluation différentes ou pour faire émerger les points essentiels d'un domaine à étudier.

2.2.2.3. Variété des différents aspects abordés et points essentiels dégagés

La synergie du *focus group* est connue pour produire une plus large variété d'informations, d'opinions et d'idées que ne le ferait l'accumulation de réponses d'individus interviewés en face à face (Merton *et al.*, 1956; Stewart et Shamdasani, 1990). Cependant, d'autres auteurs (Agar & MacDonald, 1995; Greenbaum, 2003) mettent davantage en avant le fait que le caractère synergique d'un *focus group* permet aux points clefs ou pistes à creuser de faire surface sans en donner de descriptions détaillées.

Ainsi, après avoir estimé si le *focus group* fournit plus ou moins de données détaillées que les entretiens individuels, nous proposons d'observer s'il permet d'aborder une plus large variété d'aspects d'un thème donné et/ou s'il fait davantage ressortir les points essentiels par rapport aux entretiens individuels.

Jusqu'à présent nous avons traité les différentes questions relatives à la nature du contenu produit par chaque technique. Il est aussi nécessaire de savoir quelle est la meilleure technique du point de vue de l'exploitabilité de ses données, puisque des données trop difficiles à interpréter sont susceptibles d'être tout simplement écartées.

2.2.3. Difficulté d'exploitation des données

Malgré les critiques (Morgan, 1997) adressées à leur étude ethnographique, Agar et MacDonald (1995) disent recueillir des informations qu'ils n'auraient pas comprises sans une solide connaissance du domaine forgée grâce à des études précédentes à partir d'entretiens individuels. Les participants de l'unique groupe de cette étude interagissent entre eux avec des modèles de représentations partagées et se réfèrent à des éléments implicites sans les expliquer. Ces éléments pourraient a priori être élucidés dans un questionnement de cheminement en face à face.

Nous proposons donc de chercher à savoir si l'une ou l'autre technique permet d'obtenir davantage de données faciles à interpréter.

Après avoir examiné l'ensemble des questions relatives à la nature et à l'exploitabilité des données produites, nous abordons celle de la productivité. En effet, il est aussi nécessaire de savoir laquelle des techniques génère ces données avec le plus d'efficacité.

2.2.4. Productivité en termes de temps et de coûts dans l'une et l'autre technique

En sciences sociales, en marketing et en ergonomie, le *focus group* est généralement accepté comme permettant d'obtenir une grande concentration de données avec une plus grande efficacité dans le temps (Caplan, 1990; Crabtree *et al.*, 1993; Langford & McDonagh, 2003a; Macbeth *et al.*, 2000; Stewart & Shamdasani, 1990), même si certains soutiennent que de plus amples recherches sont nécessaires pour renseigner la question (Morgan, 1997; Morgan & Krueger, 1993). D'autres données viennent aussi nuancer ces propos : Griffin et Hauser (1993) montrent qu'une heure de *focus group* est environ équivalente à une heure d'entretien individuel pour la génération de besoins utilisateurs/clients. Les résultats de Fern (1982) établissent que huit entretiens individuels produisent davantage d'idées qu'un *focus group* de huit participants et, de même, si quatre participants sont prévus pour chaque situation d'entretien. En revanche, si l'on considère les durées d'entretien (une demi heure

pour chaque entretien individuel et deux heures pour chaque *focus group*), une heure de *focus group* est à peu près équivalente à une heure d'entretien individuel⁷².

Afin de comparer correctement les deux techniques, il faut donc tenir compte de la durée des entretiens et même des diverses ressources nécessaires à la réalisation des entretiens individuels et *focus groups*.

Toutefois, il y a peu d'études systématiques de comparaison de l'une et de l'autre technique quant au temps nécessaire et aux coûts engagés pour les mener à bien (Crabtree *et al.*, 1993). Les *focus groups* seraient longs à préparer mais efficaces en concentration de données par rapport à la durée d'entretien (Crabtree *et al.*, 1993; Langford & McDonagh, 2003a). Quant à leur coût, les avis divergent. Selon Crabtree *et al.* (1993), ils seraient souvent plus onéreux que les entretiens individuels tandis que selon Morgan et Krueger (1993), ils seraient réputés peu coûteux⁷³. Il ne faut pas non plus négliger le temps et l'effort nécessaires pour préparer la recherche, pour recruter les participants, pour analyser les données et rédiger les rapports de synthèse (Langford & McDonagh, 2003a).

Nous nous proposons d'étudier la productivité des deux techniques par rapport aux coûts financiers et aux ressources humaines nécessaires à la réalisation de chaque technique, en considérant les diverses activités évoquées ci-dessus.

Après avoir exploré les diverses questions relatives à la production et à la productivité des deux techniques, nous abordons la question du nombre de sujets avant que la productivité des techniques ne stagne ou ne diminue.

2.3. LA QUESTION DU NOMBRE DE SUJETS

En général, dans une étude, que l'objectif soit de générer un maximum d'idées ou de recueillir diverses données d'évaluation, il est toujours réalisé plusieurs passations (entretiens individuels, *focus groups*, tests utilisateurs, etc.) avec des échantillons de la même population de sujets ou de différentes populations. La question est généralement de savoir quel est l'apport d'un entretien/test supplémentaire par rapport aux précédents et à partir de quel

⁷² Une relation d'équivalence peut être trouvée d'après leurs résultats entre deux *focus groups* (FG) et dix entretiens individuels (EI) (Morgan, 1997) concernant la production du nombre d'idées toutes différentes (mais non concernant le précédent résultat sur la production du nombre de meilleures idées). Donc 2 FG (2h) = 10 EI (½ h). Ainsi, 1h de FG \approx 1 ¼ h d'EI. Griffin et Hauser (1993) rapportent la même conclusion à partir d'un autre calcul.

⁷³ Mais ces auteurs émettent des désaccords concernant ce point.

nombre, il est moins utile d'en accomplir un de plus. Nous avons vu que les conclusions avancées en ce qui concerne les tests utilisateurs évoluent. La fameuse prédication « cinq tests utilisateurs suffisent ! » (Nielsen & Landauer, 1993) semble évoluer vers la nécessité d'en réaliser 10, voire 20 pour détecter un nombre suffisant de problèmes d'utilisabilité (Faulkner, 2003). Dans le domaine du marketing, des études ont cherché à répondre aux mêmes questions et mettent au jour qu'avec 20 à 30 entretiens individuels, il est possible de révéler environ 85% à 90% des besoins utilisateurs/clients (Griffin & Hauser, 1993).

Dans le même ordre d'idées, nous proposons d'analyser à partir de quel nombre d'entretiens cumulés, c'est-à-dire quel nombre de sujets, la production de données devient trop faible par rapport aux coûts engagés. Autrement dit, nous observerons le seuil à partir duquel la productivité du contenu stagne ou même diminue.

Enfin, la dernière question de recherche porte, non plus sur la comparaison de l'entretien individuel et du *focus group*, mais sur celle des situations de verbalisation avec ou sans utilisation du produit/service.

2.4. COMPARAISON D'UNE SITUATION DE VERBALISATION AVEC VERSUS SANS UTILISATION DU PRODUIT/SERVICE

En matière d'évaluation de l'utilisabilité, nous savons qu'il existe un écart entre ce que les gens disent vouloir et ce dont ils ont réellement besoin (Langford & McDonagh, 2003a; Nielsen, 1993, 1997). Il existe donc une différence entre les méthodes indirectes (entretiens, *focus groups*, etc.) et les méthodes directes où les sujets sont placés en situation d'utilisation du produit/système (tests utilisateurs, évaluations par *think aloud*, etc.).

Même s'il ne s'agit pas d'un objectif de détection de problèmes d'utilisabilité mais d'élicitation d'idées, nous proposons de même d'analyser l'impact d'une situation avec et sans utilisation du produit sur le contenu verbal produit. Comme nous le verrons au chapitre suivant, nous proposons de comparer les verbalisations obtenues par la méthode des verbalisations libres, donc avec utilisation de produits/services, et celles obtenues par la technique d'élicitation associative conjuguée à l'entretien d'explicitation, donc sans activité d'utilisation. La comparaison proposée est analogue à celle des verbalisations simultanées comparées à des verbalisations consécutives ou plutôt « ultra-consécutives » puisqu'il s'agit alors de faire verbaliser le sujet à partir du souvenir d'une (ou de plusieurs) expérience(s) vécue(s) et non pas juste après le déroulement de son action.

Les précédentes questions de recherche concernent la production de données qualitatives utiles à l'élaboration d'une évaluation quantitative. Cela correspond au premier versant de l'évaluation ergonomique que nous nous sommes proposé d'examiner. Dans ce qui suit, nous examinons le second versant, celui de l'utilisation des résultats ergonomiques (ou de résultats relatifs à d'autres aspects).

3. D'UNE LOGIQUE DE PRODUCTION DE RESULTATS D'EVALUATION A UNE LOGIQUE D'UTILISATION DES RESULTATS

3.1. ACTIVITE DE CONCEPTION ET ACTIVITE D'USAGE : DEUX AXES DE RECHERCHE COMPLEMENTAIRES

Les précédentes questions correspondent aux interrogations des professionnels du service d'évaluation. Leur souci premier est d'extraire le contenu nécessaire à la réalisation du questionnaire auprès de futurs répondants, les utilisateurs des produits/services à évaluer. La demande concernant ces évaluations, formulée par les chefs de projet des produits/services étant peu précisée (il leur est demandé d'évaluer les habitudes d'usages et la perception générale des utilisateurs/clients), les professionnels de l'évaluation cherchent à connaître les divers critères ou dimensions sur lesquels reposent le jugement des utilisateurs/clients. Ils se basent exclusivement sur cette approche⁷⁴. Ils tiennent compte de la relation aux artefacts seulement du point de vue de l'activité des utilisateurs.

Or, s'il s'agit de traiter la relation aux artefacts comme problème central, cette dernière doit être explorée selon deux axes correspondant à deux formes d'activités humaines qui ne peuvent être pensées indépendamment : celles des utilisateurs des produits/services, considérés avec leur culture et histoire dans leur contexte singulier et celles des divers acteurs de la conception (Bannon & Bodker, 1991). En effet, l'ergonome doit s'intéresser de plus près à l'activité des concepteurs pour fournir « *des aides réelles aux concepteurs* » (Rabardel, 1995, p. 50) ou encore des « *préconisations en ne perdant pas de vue qu'il s'agit là d'un outil de travail pour le concepteur* » (Roussel, 1996, p. 111). Smit (1996, p. 20) résume le problème ici posé de la façon suivante : « *We always preach that the conceptual model of an application must be logical and clear for the users, to enable them to build a correct mental*

⁷⁴ Divers facteurs relatifs au contexte général de l'entreprise interviennent aussi dans ce parti pris et sont décrits au chapitre 4.

model. We must apply the same principles to clarify our usability approach to people with an information technology based mental model ».

Dans l'étude présentée au paragraphe 2, nous cherchons à produire des connaissances qui aident les professionnels du service d'évaluation à mieux choisir les deux techniques comparées (voire les autres techniques d'élicitation d'idées et pensées mobilisées) en lien avec leurs conditions de réalisation (ex. leur niveau de connaissance d'un domaine de produits/services et de leurs utilisateurs/clients). Cependant, dans cette étude, nous ne tenons compte que de l'activité des utilisateurs des produits/services et de celle des évaluateurs, mais non de l'activité des divers acteurs de la conception, c'est-à-dire des destinataires des résultats d'évaluation qui les utilisent pour transformer les produits/services.

Aussi dans un second temps, nous proposons de les considérer. Nous avons déjà évoqué les chefs de projets, demandeurs des évaluations. Ils sont les premiers destinataires des résultats d'évaluation, mais ils ne sont pas les seuls. Les professionnels du marketing et de la conception/réalisation interviennent également pour faire évoluer les produits/services sur la base des résultats d'évaluation, qu'ils y aient accès directement ou non. Ils peuvent avoir accès seulement à la reformulation des chefs de projet suite à leurs prises de décision à partir des résultats (cela équivaut à l'écart entre maîtrise d'ouvrage et maîtrise d'œuvre). En prenant en considération ces acteurs, il s'agit de chercher à savoir quelles sont les connaissances opérationnelles, c'est-à-dire les résultats d'évaluation qui leurs sont utiles pour pouvoir effectivement transformer et améliorer les produits/services. L'enjeu est, non pas simplement de leur apporter des connaissances sur les utilisateurs/clients telles des « substances essentielles » à incorporer dans leur conception des produits/services, mais aussi de savoir quelle est la réalité de leur activité, leurs réelles marges de manœuvre (ex. nécessité d'arbitrage entre différentes contraintes ; Baber & Stanton, 1996; Chevalier & Ivory, 2003; Roussel, 1996), le type de résultats dont ils peuvent se servir et ceux dont ils ne peuvent pas (ex. en raison de leur imprécision ; cf. chapitre 4) et leurs représentations et compréhensions liées à leurs savoir-faire et à leurs connaissances spécifiques (ex. Roussel, 1996).

Il s'agit, comme le propose Smit (1996) dans sa méthodologie d'évaluation, de partir de la situation des destinataires et de chercher à savoir de quels résultats ils ont besoin (sans pour autant supprimer tout contenu ergonomique). La question peut aussi se formuler de la façon suivante : quels résultats d'évaluation sont susceptibles d'être appropriés par les divers acteurs de la conception et aussi du marketing ? L'idée est de considérer les professionnels de

la transformation en tant qu'utilisateurs des résultats d'évaluation⁷⁵ et ces derniers comme

artefacts, auxquels les sujets sont susceptibles d'attribuer des fonctions (ex. permettant une prise de décisions par rapport à la transformation du produit/système). Les sujets sont donc susceptibles d'instituer ou non ces artefacts comme moyens de réalisation de leur activité.

Ce travail s'inscrit dans l'approche instrumentale développée par Rabardel (1995) dans laquelle les artefacts sont considérés en référence à l'activité humaine. Nous présentons cette approche dans ce qui suit.

3.2. L'APPROCHE INSTRUMENTALE

L'instrument est une entité mixte. Il ne se réduit pas à l'objet, l'artefact⁷⁶ matériel ou symbolique. Il comprend aussi une composante liée au sujet : ce dernier élabore les schèmes d'utilisation pour s'approprier et inscrire l'artefact en tant que moyen dans la réalisation de son activité.

Les schèmes d'utilisation constituent des organisateurs de l'action et se décomposent en deux sortes : les schèmes d'usage et les schèmes d'action instrumentée. Les premiers, les plus élémentaires, se rapportent aux tâches secondaires, c'est-à-dire à la gestion des caractéristiques particulières de l'artefact (ex. schème de lecture des graduations des axes des abscisses ou des ordonnées d'un graphique, schème d'identification de la valeur d'une mesure à partir de la hauteur de l'histogramme selon les graduations du graphique, etc.). Les seconds, constitués d'une séquence de schèmes élémentaires, sont relatifs aux tâches principales et concernent l'objet de l'activité (ex. schème de déduction des points faibles du produit/système à partir d'un graphique par histogramme rapportant son évaluation subjective multidimensionnelle).

De plus, les schèmes sont sociaux. Ils sont appropriés et développés par un individu au contact des autres. Ils sont transmis d'une génération à une autre au sein d'un groupe d'individus. Les schèmes peuvent aussi évoluer d'une génération ou d'une époque à une autre. Avec les nouvelles technologies, par exemple, certains processus perceptifs, cognitifs et communicatifs se transforment. Les travaux de Nosulenko et Samoylenko (2001) révèlent que les processus ou schèmes d'écoute des jeunes sont altérés par les technologies d'enregistrement sonore. « *Pour eux le son [d'instruments musicaux] le plus naturel sera le son altéré par le matériel électroacoustique* » (Nosulenko & Samoylenko, 2001, p.4). Ainsi, les instruments médiatisent le rapport du sujet au monde. Ces médiations peuvent être de trois natures :

- la médiation pragmatique : l'instrument permet d'opérer une action de transformation de l'objet sur lequel porte l'activité,
- la médiation épistémique : l'instrument assure une fonction de développement de connaissances quant à l'objet de l'activité,
- la médiation réflexive : l'instrument est un moyen de contrôle et de régulation de l'activité.

Cependant les systèmes techniques, outils de travail, objets du quotidien ou tout autre objet ne sont pas d'emblée des instruments pour les hommes. Ils le deviennent au cours de processus de genèse instrumentale. Ces processus peuvent être ceux de l'instrumentation qui concernent le développement des schèmes d'utilisation déjà évoqués et qui se réalisent au cours de l'activité du sujet. Ces processus peuvent aussi être ceux de l'instrumentalisation qui concernent l'artefact. Et enfin, le plus souvent, les deux types de processus se développent conjointement :

- dans le processus d'instrumentation, les schèmes sont constitués et/ou évoluent (par accommodation, coordination, combinaison de schèmes, etc.) ou bien des artefacts nouveaux sont assimilés à des schèmes existants,
- le processus d'instrumentalisation est un processus de découverte et d'attribution de fonctions (pragmatique, épistémique, réflexive) aux objets ainsi transformés en instruments. L'artefact peut aussi connaître une transformation au niveau de ses composants (ajout, suppression, sélection, regroupement, etc.).

L'approche instrumentale préconise de concevoir tout artefact en développant une activité itérative d'étude du développement de ces processus et de conception à partir de cette connaissance.

C'est l'approche que nous adoptons⁷⁷ pour concevoir un artefact instrumentalisable par les professionnels de la transformation du Web marchand, pour les outiller dans leur activité d'identification de leurs besoins en connaissances, c'est-à-dire ce qu'ils ont besoin de savoir à l'issue d'une évaluation et dans leur activité d'élaboration de cette évaluation.

Dans ce qui suit, nous examinons brièvement les ressources disponibles pour l'activité de conception et de marketing des sites Internet marchands : les guides de recommandations et les études d'évaluation et surtout si ces dernières prennent en compte l'activité des professionnels du Web marchand.

3.3. LA PRISE EN COMPTE DE L'ACTIVITE DES PROFESSIONNELS DU WEB MARCHAND DANS L'EVALUATION

Il existe de nombreux travaux concernant la qualité perçue et les problèmes d'utilisabilité des sites Internet marchands, mais très peu prennent en compte et étudient de près l'activité des destinataires et leurs besoins en connaissances.

3.3.1. Des évaluations, des guides de recommandations mais peu d'études sur les « besoins » des professionnels de la transformation

Il existe principalement des études exploratoires pour définir qui sont les cyber-consommateurs et les internautes (Bellman *et al.*, 1999; Katz & Aspden, 1997; Kehoe *et al.*, 1997; Kehoe *et al.*, 1998a; Kehoe *et al.*, 1998b; Lohse *et al.*, 2000; Pitkow & Kehoe, 1996), d'une part, et, d'autre part, des études multidisciplinaires sur les facteurs ergonomiques, « sensoriels » et commerciaux influençant la qualité perçue des internautes et cyber-consommateurs (une synthèse est proposée par Helander & Khalid, 2000).

Ces dernières sont, par exemple, réalisées :

- par questionnaire en ligne (Lee *et al.*, 2000; Teo, 2002),
- par questionnaire « prêt à l'emploi » ou construit aux fins de l'étude, avec des tâches d'utilisation préalables de sites marchands (Dellaert & Kahn, 1999; Jarvenpaa *et al.*, 2000; Kim & Lee, 2002; Kim & Yoo, 2000; Konradt *et al.*, 2003; Kwon *et al.*, 2002; Leulier *et al.*, 1999; Spool *et al.*, 1999; Teo *et al.*, 2003; Yoo & Kim, 2000) et avec d'éventuels *focus groups* à la suite (Jarvenpaa & Todd, 1997),
- en utilisant la *repertory grid* (Hassenzahl & Trautmann, 2001),
- par entretiens approfondis avec des tâches d'utilisation préalables (Lindgaard & Dudek, 2002),
- par la technique des protocoles verbaux avec des scénarios de tâches définis (Nielsen *et al.*, 2001; Tilson *et al.*, 1998) ou une liberté complète donnée à l'utilisateur de parcourir le site et même l'absence de l'expérimentateur (Rehman, 2000),
- par tests de performance (Bouch *et al.*, 2000; Pilverdier & Gaillard, 2001).

Enfin, on peut citer une étude qui interroge des Webmasters de sites Internet de divers secteurs industriels. Cette étude recueille cependant des données sur les facteurs influençant la perception des internautes et non directement sur les besoins en connaissances des Webmasters, même si par cette démarche, elle s'en approche.

Les diverses connaissances constituées à partir de ces études ou celles développées par des praticiens lors de leur activité de conception et de marketing ont aussi permis d'établir de nombreux guides :

- expliquant le déroulement d'un processus de conception de sites Internet (Goto & Cotler, 2002),
- comprenant des recommandations de conception des sites Internet (Apple, 1996; IBM, 1999; Leulier *et al.*, 1998; Nielsen, 2000; Nogier, 2002; Shneiderman, 1997; Spool *et al.*, 1999; W3C, 1999),
- et/ou comprenant des recommandations concernant les technologies et stratégies commerciales pour développer une activité marchande par Internet (Fouchard, 2001; Ghernaouti-Hélie, 2002; Jarvenpaa & Todd, 1997; Lohse & Spiller, 1998a, 1998b; Macarez & Leslé, 2001).

Ces guides fournissent des renseignements précieux et opérationnels pour les professionnels de la transformation. Cependant, à notre connaissance, il n'existe pas de travaux qui formalisent l'ensemble de leurs besoins en connaissances ou qui analysent de près leur activité d'utilisation de résultats d'évaluation ou d'éléments de recommandations dans le domaine du Web marchand.

3.3.2. Les questions de recherche dans ce domaine

Ainsi, nous cherchons :

- à identifier leurs besoins en connaissances pour piloter les actions commerciales et la conception de leurs produits/services,
- et plus généralement à comprendre les exigences de leur activité.

Ce terme « besoin », emprunté au domaine du marketing, ne correspond pas à l'idée, in fine, de n'observer que des dimensions relatives aux stratégies commerciales et non plus à tout autre aspect, y compris ergonomique, mais bien à l'optique d'outiller les professionnels du marketing et de la conception et, en ce sens, de s'interroger sur comment ils peuvent l'être. Quels sont leurs besoins en termes d'outillage ou de connaissances ? Quels sont leurs représentations, leurs concepts et leur vocabulaire ? Comment sont-ils susceptibles de comprendre divers résultats ? Comment et à quoi sont utilisés des résultats d'évaluation potentiels ? Quelles sont les caractéristiques précises des sites marchands qu'ils peuvent remettre en question ? Concernant quels éléments relatifs à toute l'activité de commerce par

Internet, y compris l'interface, doivent-ils prendre des décisions en termes de modification ? Quels arbitrages opèrent-ils entre contraintes ? etc.

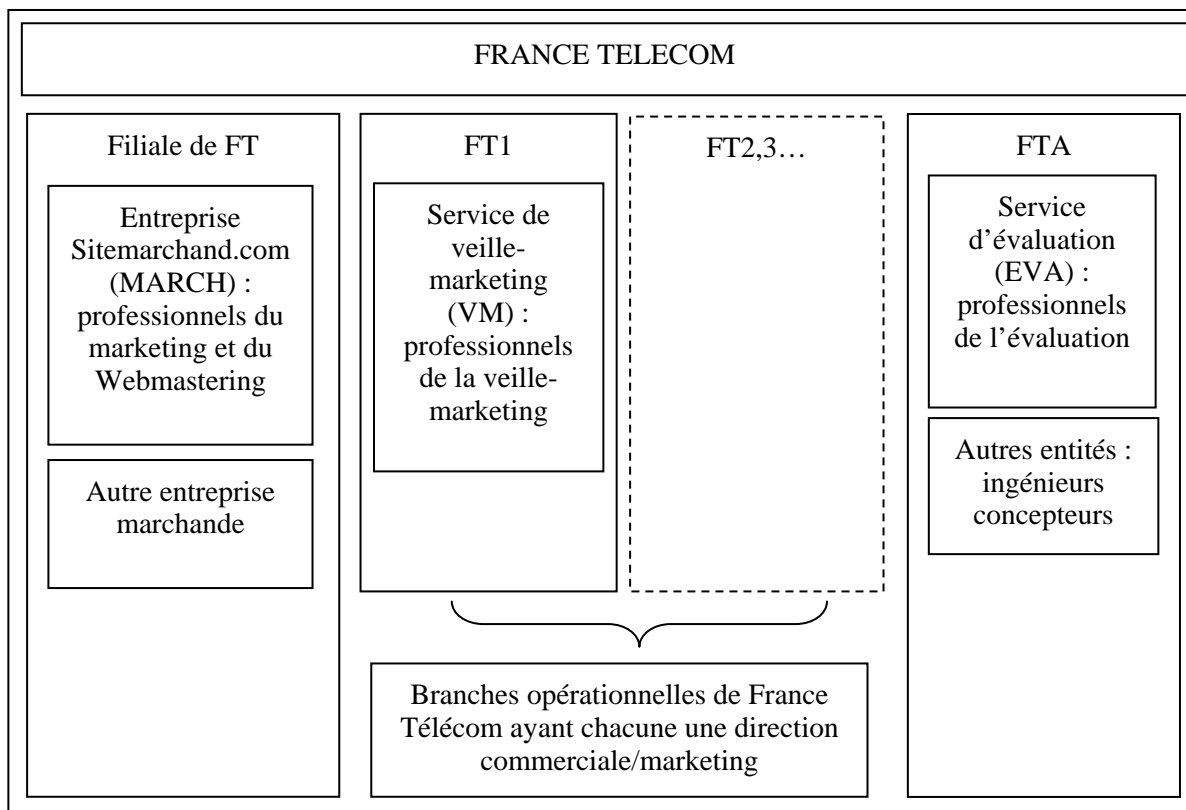
4. LES TERRAINS ET LA DEMARCHE DE RECHERCHE

Nous présentons maintenant les deux terrains de recherche, à savoir, d'une part, le service d'évaluation où la première partie de la thèse s'est réalisée et d'autre part, le service de veille-marketing et l'entreprise marchande par Internet à partir desquels s'est déroulée la recherche de la seconde partie de la thèse. Différentes branches et filiales forment le groupe France Télécom dont nous exposons une partie de son organisation.

4.1. L'ORGANISATION DE FRANCE TELECOM ET LES ENTITES CONCERNEES PAR LA RECHERCHE

Pour des raisons de confidentialité, nous ne divulguons pas la réelle dénomination des entités de France Télécom. Le service d'évaluation fait partie de la branche appelée FTA de France Télécom, tandis que le service de veille-marketing est une entité d'une des branches opérationnelles de France Télécom appelée FT1 et enfin l'entreprise marchande par Internet (Sitemarchand.com) est une des entreprises d'une filiale de France Télécom. Ces différents éléments sont récapitulés dans l'organigramme simplifié (cf. figure 1 ci-après).

Figure 1 : Organigramme simplifié des branches et filiales de France Télécom concernées dans cette thèse



Nous décrivons succinctement les différentes entités concernées par cette thèse et leurs éventuelles interactions.

4.1.1. La branche FTA de France Télécom

- Le service d'évaluation (EVA) et ses professionnels de l'évaluation :

Il s'agit d'un service exclusivement consacré à une activité d'évaluation des produits/services développés par les autres services de FTA, les branches opérationnelles ou la filiale de France Télécom (cf. organigramme ci-dessus).

Ce service assure une activité d'évaluation des produits/services :

- parfois au cours de développement des produits/services,
- très souvent à la fin du processus de conception, une fois le produit/service opérationnel, pour vérifier qu'il ne présente pas de défauts majeurs avant son lancement sur le marché,
- ou encore une fois le produit/service mis sur le marché : il s'agit alors d'études exploratoires, généralement de type « benchmark », en phase de « feed-back ». Une

méthodologie d'évaluation a été élaborée regroupant diverses techniques d'évaluation à réaliser de façon itérative (évaluation quantitative par questionnaire, analyses de remontées clientèles⁷⁸, évaluations psychophysiques, mesures objectives itératives de paramètres physiques, etc.). L'idée est de fournir régulièrement des indicateurs quant aux habitudes d'usage et à l'appréciation subjective qualitative d'une famille de produits (ex. les sites Internet portail de France Télécom et leurs concurrents, les sites Internet marchands de France Télécom et leurs concurrents, etc.) éventuellement corrélés à des mesures objectives, pour guider les chefs de projet des produits/services dans leurs prises de décision. Si nécessaire, les dimensions d'évaluation identifiées au départ doivent être renouvelées.

Dans cette thèse, nous nous intéressons à la dernière catégorie d'évaluations.

Les professionnels de ce service sont issus d'une formation en ergonomie, en psycho-acoustique, en linguistique ou en statistique et sont isolés des acteurs et de l'activité de marketing et de conception. Ils sont seulement en interaction avec les demandeurs des études, généralement les chefs de projet (issus d'une formation en marketing ou en ingénierie), et avec certains professionnels de la transformation (marketing/conception), auxquels ils présentent les résultats et conclusions des études.

- Les autres entités de la branche FTA et ses ingénieurs :

De très nombreuses entités existent dans la branche FTA dont certaines assurent une activité de conception. Les ingénieurs de ces entités sont fréquemment sollicités par les branches opérationnelles ou filiales de France Télécom pour développer des technologies et concevoir leurs produits. Par exemple, le responsable du Webmastering de l'entreprise marchande par Internet de la filiale de France Télécom (détaillée au paragraphe 4.1.3. ci-dessous) a supervisé un projet de reconception de la recherche (interne) par mots-clefs sur leur site Internet marchand dont la programmation informatique était remaniée par des ingénieurs de FTA.

4.1.2. La branche opérationnelle FT1 et son service de veille-marketing (VM)

Un des services de la branche opérationnelle nommée FT1 est dédié à une activité de veille-marketing concernant des produits/services développés dans les différentes branches opérationnelles et, en particulier, dans la filiale de France Télécom. Les professionnels de ce

service développent une activité de marketing stratégique (études de marché, études comparatives des offres en termes de produits, enquêtes par questionnaire, mesures d'audience, etc.) portant sur les activités par Internet et nouvelles technologies de France Télécom et notamment sur l'un des sites marchands de la filiale de France Télécom, détaillée ci-dessous. Ils assurent une activité de conseil, entre autres, auprès de l'entreprise Sitemarchand.com.

4.1.3. La filiale de France Télécom, l'entreprise Sitemarchand.com (MARCH) et ses professionnels du Web marchand

Différentes entreprises marchandes par Internet font partie de la filiale de France Télécom. Nous avons pu rencontrer des professionnels de la conception et du marketing d'une de ces entreprises que nous nommons Sitemarchand.com. Cette entreprise assure une activité de conception/maintenance de son site Internet, de prospection et de vente de produits en ligne, d'animation commerciale et publicitaire, de stockage et distribution de ses produits et enfin de service clientèle par sa « hotline ». Elle comprend donc un service produits, un service logistique, un service clients, un service informatique, une équipe Webmastering et une équipe marketing (comprenant elle-même diverses cellules : marketing de fidélisation, recrutement clients, communication, etc.). Nous présentons ci-dessous les trois professionnels que nous avons rencontrés.

- Le responsable du Webmastering du Sitemarchand.com :

Le service Webmastering est l'équipe qui assure la conception des pages HTML⁷⁹ du Sitemarchand.com et, quand nécessaire, la mise en forme d'un questionnaire d'évaluation diffusé par Internet⁸⁰. L'activité du responsable du service Webmastering de l'entreprise Sitemarchand.com concerne la spécification de la conception du site Internet (ex. architecture du site, principes de navigation, procédures de commande sur le site, etc.) et l'encadrement de son équipe, à savoir les Webdesigners et le responsable de la charte graphique.

- Un autre professionnel du Webmastering, le responsable de la charte graphique :

Le responsable de la charte graphique, comme son titre l'indique, veille à la mise en place et au respect d'une charte graphique (choix de couleurs, de polices de caractères, etc.) dans la conception des pages HTML du site Internet marchand. Il réalise également des études graphiques pour améliorer l'identité visuelle de l'enseigne de l'entreprise, véhiculée par le site Internet et les divers supports publicitaires.

- Un professionnel de l'équipe marketing, chargé des études d'évaluations et du marketing de fidélisation :

Le marketing de fidélisation est une activité de développement de stratégies de communication et d'offres commerciales dont le but est de conserver les personnes déjà clientes du Sitemarchand.com. Le professionnel assurant cette activité est également le chargé des études d'évaluation. Il s'occupe de rassembler les différents « besoins en évaluation » ou plutôt divers problèmes identifiés émanant de tous les services de l'entreprise (marketing, Webmastering, logistique, produits, clients), de décider avec le directeur marketing des objectifs d'évaluation prioritaires et de construire le questionnaire en partenariat avec l'un des professionnels du service de veille-marketing (appelé le professionnel de la veille-marketing, spécialiste du Sitemarchand.com).

- Les ressources d'évaluation de l'entreprise Sitemarchand.com :

L'entreprise Sitemarchand.com a ses propres ressources d'évaluation. Des enquêtes par questionnaire auprès de différents segments d'utilisateurs/clients sont régulièrement organisées (au moins une par an). Le professionnel du marketing chargé des études de même que le professionnel de la veille-marketing spécialiste du Sitemarchand.com sont donc aussi des acteurs de l'évaluation mais doivent être distingués des professionnels de l'évaluation (du service d'évaluation de FTA) dont l'activité est exclusivement dédiée aux évaluations de produits/services développés par d'autres services et branches opérationnelles.

Par ailleurs, les professionnels du Sitemarchand.com commandent parfois des évaluations à d'autres services. Le responsable du Webmastering a occasionnellement commandé des évaluations relatives à la perception de l'interface auprès du service de veille-marketing de FT1. Une évaluation subjective relative à la recherche (interne) par mots-clefs sur le site Sitemarchand.com a aussi été commandée au service d'évaluation de FTA. Mise à part cette étude, les professionnels de l'évaluation n'ont pas été sollicités par l'entreprise Sitemarchand.com.

4.2. LES ACTIVITES DE RECHERCHE AU SEIN DES DEUX TERRAINS

Dans ce qui suit, nous replaçons les activités correspondant aux deux mouvements développés dans cette thèse par rapport aux deux terrains de recherche. L'organisation de la suite du texte est également esquissée dans la conclusion.

4.2.1. Les activités sur le premier terrain de recherche

Le service d'évaluation de la branche FTA nous a accueilli durant trois ans dans le cadre d'un contrat à durée déterminé de thèse. A travers nos travaux dans ce service et lors de réunions de travail ou de discussions informelles, nous avons appris à connaître l'un des processus d'évaluation qu'ils ont élaboré ainsi que leurs interrogations relatives à leurs méthodes d'évaluation. Nous avons ainsi mené l'étude comparative des deux techniques, utilisées pour l'élaboration de dimensions subjectives et de contenu verbal, en vue de la construction d'un questionnaire d'évaluation dans le domaine du Web marchand.

4.2.2. Les activités sur le second terrain de recherche

Le second terrain correspond au mouvement de renversement du point de vue porté sur les résultats d'évaluation. Ici, la démarche est à la fois celle d'une recherche et analyse des besoins en connaissances (« besoins en évaluation ») de professionnels de la transformation et celle d'une démarche d'engineering et de conception d'un artefact. Après la précédente étude, notre démarche devient très empirique, « non méthodologique » et suit un processus historique peu courant dans une thèse mais qui est ici nécessaire étant donné nos contraintes.

4.2.2.1. Des difficultés qui occasionnent une démarche « très empirique »

Des raisons liées à notre statut de salarié du service d'évaluation ne nous ont pas permis de rencontrer rapidement tous les interlocuteurs souhaités. C'est à partir d'un des professionnels de la veille-marketing que nous avons pu de « fil en aiguille » prendre contact avec d'autres professionnels du marketing et de la conception de sites Internet marchands.

Ainsi, il n'a pas été possible d'analyser de façon approfondie l'activité d'utilisation de résultats d'évaluation en suivant les différentes itérations d'évaluation et de conception d'un produit/système précis (ex. l'utilisation de résultats issus d'une évaluation par questionnaire auprès d'utilisateurs d'un site Internet marchand et l'activité de transformation à partir de ces résultats). Nous avons cependant pu mener des entretiens auprès des différents professionnels rencontrés au fur et à mesure.

A défaut de résultats d'évaluation, nous sommes parti des dimensions utilisateurs de sites Internet marchands, issues de la précédente étude, que nous avons présentées à divers professionnels (voir la description des entretiens menés plus bas). Nous avons procédé à une première transformation de surface de cette liste de dimensions pour en améliorer sa compréhension, puis à son opérationnalisation en profondeur pour concevoir un véritable outil d'aide la conception d'évaluations. Cette partie suit donc aussi un processus chronologique car elle décrit la transformation puis la conception d'un artefact qui sont, par définition, des processus progressifs et itératifs.

4.2.2.2. Récapitulatif des entretiens menés

Nous avons mené différents types d'entretiens ou exercices (ex. hiérarchisation des dimensions utilisateurs présentées), au fur et à mesure des professionnels rencontrés et selon le besoin d'approfondir certains points ou de valider de nouveaux artefacts conçus. Nous les récapitulons dans le tableau suivant (sans les décrire en détail, car ils le sont aux chapitres 4 et 6 ; cf. tableau 1 ci-après) qui constitue un repère pour la lecture de l'ensemble du texte.

Tableau 1 : Récapitulatif des entretiens menés avec les professionnels du marketing et de la conception et des actions de transformation ou de conception d'artefacts destinés à ces derniers

	1 ^{er} prof. ^{el} VM - spécialiste MARCH	2 ^{ème} prof. ^{el} VM	2 ingénieurs FTA	resp. Webm. MARCH	resp. ch. graph. MARCH	mktg fidél. MARCH	N° Cha- pitre
Type de sujet	1 et 2	2	2	2	2	1 et 2	
Entretiens libres	✓	✓					4
Exercice de hiérarchisation des dimensions et sous-dimensions	✓	✓					4
Entretiens avec utilisation du Sitemarchand	✓	✓	✓				4
Transformation de surface de la liste exhaustive de dimensions	Réalisée à partir des précédents entretiens avec les deux professionnels de la veille-marketing et avec les deux concepteurs multimédia.						4
Entretien de validation de la transformation				✓			4
Ebauche d'un outil d'élaboration d'évaluations	Réalisée à partir d'une seconde analyse plus approfondie de tous les précédents entretiens aboutissant à l'insuffisance de la transformation de surface de la liste de dimensions et du besoin d'un outil opérationnel d'aide à l'élaboration d'évaluations.						5
Entretiens de validation de l'ébauche	✓			✓	✓	✓	6
Outil final	Obtenu après les modifications suggérées dans les entretiens de validation ci-dessus.						6

Légende : VM : veille-marketing ; MARCH : entreprise Sitemarchand.com ; Webm. : service Webmastering ; ch. graph.: charte graphique; mktg. fidél. : marketing de fidélisation ; sujets type 1 : ont une activité d'évaluation ; sujets type 2 : ont une activité de conception /marketing des produits/systèmes

4.2.3. Trois types de sujets

Dans le précédent tableau, nous avons introduit la notion de sujets de type 1 et 2 que nous expliquons ci-dessous. Dans cette thèse, nous considérons, en effet, et faisons référence à trois types de sujets :

- Les sujets de type 1 :

Les sujets de type 1 sont ceux qui ont une activité d'évaluation. Par exemple, on l'aura compris, les professionnels du service d'évaluation sont exclusivement des sujets de type 1 (ils ne sont pas présents dans le tableau ci-dessus).

- Les sujets de type 2 :

Les sujets de type 2 sont ceux qui ont une activité de transformation des produits/systèmes (y compris des aspects marketing) et d'utilisation des résultats d'évaluation. Certains acteurs peuvent être, à la fois, des sujets de type 1 et des sujets de type 2, comme c'est le cas du professionnel de la veille-marketing du site marchand qui, à la fois, participe aux études d'évaluation du Sitemarchand.com, puis conçoit et propose des conseils au niveau de certaines orientations de marketing stratégique du site. C'est encore plus évident avec le professionnel marketing qui est chargé de la réalisation des études d'évaluation et qui utilise les résultats produits (et d'autres données) pour élaborer des stratégies de communication et des offres commerciales pour fidéliser les clients du Sitemarchand.com. Le statut des sujets dépend donc de leur activité à un moment donné et non du poste qu'ils occupent.

- Les sujets de type 3 :

Les sujets de type 3 sont les utilisateurs des produits/services et, dans le cas de cette thèse, des « cyber-consommateurs », internautes du Web et consommateurs de produits vendus par Internet.

5. CONCLUSION

Cette thèse aborde les deux versants de l'évaluation ergonomique, à savoir la production de résultats et leur utilisation. Ce travail cherche à contribuer à la question suivante : quels résultats faut-il produire pour qu'ils soient, d'une part, valides, et d'autre part, pertinents, c'est-à-dire utilisables par les professionnels qui doivent améliorer la conception et le marketing des produits/services ? La thèse est organisée en deux mouvements :

- Le versant concernant la production de résultats ou plutôt, dans notre cas, de dimensions d'évaluation :

Le chapitre 3 relate l'étude sur la production de dimensions utilisateurs et de contenu verbal, dans le cadre d'une étude qualitative amont à une évaluation par questionnaire. Des résultats sont produits concernant la comparaison de dimensions utilisateurs et de dimensions experts, concernant la comparaison de l'entretien individuel et du *focus group* (nature, exploitabilité et productivité de leur contenu verbal), concernant le nombre de sujets avant un déclin de la productivité et concernant la comparaison des verbalisations selon qu'elles sont

produites lors d'une situation avec ou sans utilisation de produits/services. Enfin, une liste de dimensions utilisateurs de sites Internet marchands est générée.

- Le versant concernant l'utilisation des résultats ou plutôt, dans notre cas, la construction de l'outil d'aide à la production de résultats adaptés aux professionnels de la transformation :

Dans le chapitre 4, charnière de la thèse, nous recueillons les réactions des professionnels de la veille-marketing et de la conception de sites Internet (marchands) quant à la liste de dimensions utilisateurs pour constater l'impossibilité de son instrumentalisation par ces derniers, même après une transformation de surface pour en améliorer sa compréhension.

A ce stade, une activité de conception d'un outil d'aide à l'élaboration d'évaluations pour les professionnels du Web marchand (producteurs et utilisateurs de résultats d'évaluation) est entreprise.

Dans le chapitre 5, nous analysons leurs besoins en connaissances, puis construisons l'outil à partir de ces dernières. Nous cherchons également à faciliter les processus potentiels d'instrumentation et d'instrumentalisation de l'outil, à travers les modalités de formulation et de présentation de ce contenu.

Le chapitre 6 décrit la version finale de l'outil, « besoin en évaluation » par « besoin en évaluation », en rapportant les éléments de la littérature du domaine et de l'activité des professionnels qui les justifient et en nourrissent le contenu.

CHAPITRE 3

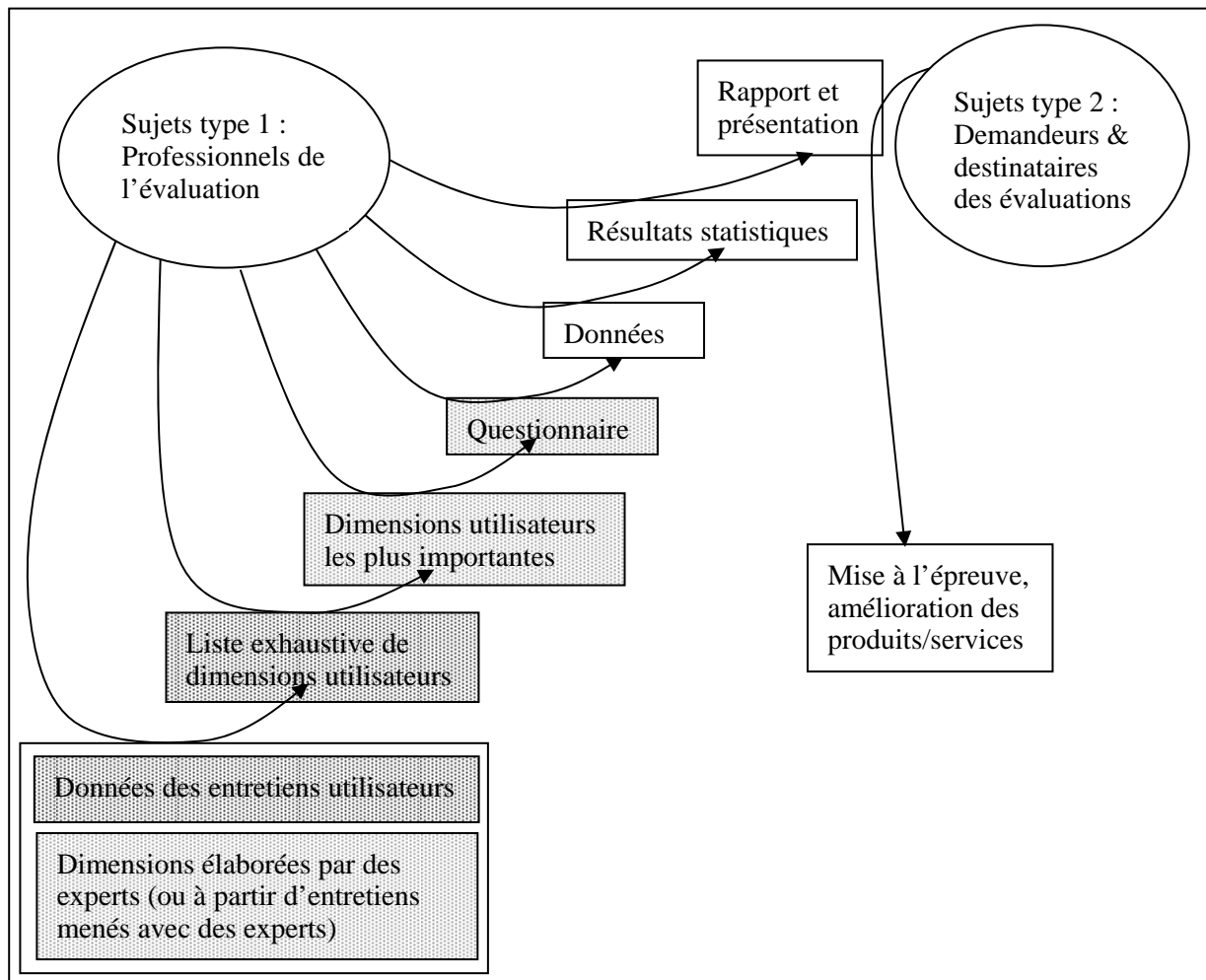
ANALYSE COMPARATIVE DE L'ENTRETIEN INDIVIDUEL ET DU *FOCUS GROUP* POUR LA CONSTRUCTION D'UN QUESTIONNAIRE D'EVALUATION

REPERES POUR LE CHAPITRE

Ci-dessous nous avons modélisé le processus d'évaluation, développé par les professionnels du service d'évaluation, qui comprend l'extraction de dimensions utilisateurs, la sélection des dimensions les plus importantes, la construction du questionnaire à partir de ces dernières, etc. L'étude comparative (entretien individuel versus *focus group*) explicitée dans ce chapitre s'inscrit dans ce même processus. Elle porte sur l'extraction de dimensions utilisateurs de sites Internet marchands (parties en gris foncé dans la figure 2). Les conclusions de l'étude concernent néanmoins les étapes suivantes jusqu'à la construction du questionnaire (parties en gris clair). Ce schéma servira aussi à situer le travail réalisé au chapitre suivant où le processus complet y est traité puisque nous abordons le renversement de point de vue porté sur l'évaluation et l'activité des différents acteurs dont les sujets de type

?

Figure 2 : Processus élaboré par les professionnels du service d'évaluation, depuis l'extraction des dimensions utilisateurs jusqu'à la restitution des résultats d'évaluation aux professionnels destinataires



Légendes : Les sujets de type 1 ont une activité d'évaluation et les sujets de type 2 ont une activité de conception/marketing des produits/systèmes.

1. INTRODUCTION

Ce chapitre correspond au premier versant traité dans cette thèse : la production de résultats d'évaluation ou plus exactement la production de dimensions utilisateurs qui précède celle de ces résultats. Les travaux exposés ici concernent la production de données subjectives et verbales pour dégager des dimensions de perception des utilisateurs des sites Internet marchands. Ces dernières servent à la construction d'une évaluation exploratoire par questionnaire, en phase de « feed-back ».

Un des objectifs de l'étude est de produire des connaissances opérationnelles qui répondent aux interrogations des professionnels de l'évaluation. Ainsi, elle s'inscrit dans un processus élaboré par ces professionnels (cf. figure 2, page précédente) en lien avec une demande d'évaluation peu spécifiée au départ, formulée par les chefs de projet de produits/services faisant partie d'autres entités de FTA, de branches opérationnelles (FT1, FT2, etc.) ou d'entreprises de la filiale de France Télécom. L'autre objectif de l'étude est d'élaborer des éléments de connaissances plus généraux qui renseignent les questions de recherche issues de l'examen de la littérature établi au chapitre précédent.

Deux techniques ont été utilisées pour faire éliciter des idées et pensées aux cyber-consommateurs : l'entretien individuel et le *focus group*. Les éléments de jugement obtenus permettent de construire des dimensions qui sont comparées à des dimensions élaborées par des experts. Nous examinons l'intérêt de données utilisateurs par rapport à des données experts, puis comparons l'entretien individuel et le *focus group* en termes de nature et d'exploitabilité des données produites.

Les deux techniques sont constituées d'une phase de verbalisation sans utilisation de sites marchands, puis avec utilisation. La première phase mobilise la technique d'élicitation associative et l'entretien d'explicitation et la seconde phase, la méthode des verbalisations libres utilisée dans l'approche systémique de la psychologie appliquée à la qualité perçue⁸¹. Nous comparons les données obtenues dans l'une et l'autre phase.

Après avoir examiné la nature et l'exploitabilité des contenus dans les diverses situations (individuelle versus collective et sans versus avec utilisation), nous cherchons à renseigner la productivité des deux techniques. Les productivités de chacune en termes de ressources humaines et financières engagées sont comparées. Enfin, l'évolution de la productivité d'un entretien à un autre est étudiée pour estimer le nombre de sujets minimum suffisant.

Les résultats nous amènent à discuter de la pertinence des deux techniques pour les professionnels de l'évaluation de France Télécom. Les conclusions sont directement valables ou élargies aux champs de l'évaluation ergonomique et des HCI où se posent les questions de la pertinence et de la productivité des deux techniques. Enfin, nous élargissons l'interprétation de certains de nos résultats à l'utilisation des techniques pour la conception.

Comme nous l'avons explicité au chapitre 1, la qualité perçue des sites Internet marchands ne se limite pas au seul champ de l'ergonomie « traditionnelle » avec les concepts d'utilité et d'utilisabilité, mais compte également des dimensions affectives, émotionnelles et mêmes commerciales. Avec les sites Internet commerciaux, « *les "utilisateurs" deviennent également des consommateurs* » et « *l'ergonomie doit s'interroger sur ses objectifs et ses méthodes* » (Leulier *et al.*, 1999, p. 49). Tout en préservant la limite donnée entre ergonomie et marketing, nous reconnaissons l'intérêt de l'évolution de l'ergonomie vers d'autres « dimensions-sœur ». Cette étude y contribue. Notamment, des dimensions relatives au marketing, comme la « notoriété perçue » ou l'impact des prix dans l'appréciation globale, sont énoncées par les cyber-consommateurs et prises en compte.

2. OBJECTIFS DE L'ETUDE

La comparaison de l'entretien individuel et du *focus group* doit se réaliser en approchant le plus possible les conditions de réalisation des professionnels de l'évaluation, puisqu'un des objectifs de l'étude est de produire des connaissances opérationnelles pour eux. Ceux-ci utilisent essentiellement l'entretien individuel et, plus récemment, de façon expérimentale, le *focus group*, dans la phase qualitative amont à la construction de l'évaluation. Les professionnels suivent tout un processus d'évaluation modélisé dans la figure 2 et détaillé, en partie, dans ce qui suit⁸². Pour une comparaison valide, notre étude s'inscrit donc également dans ce processus, même si toutes les étapes décrites n'ont pas été réalisées.

2.1. DESCRIPTION DU PROCESSUS DANS LEQUEL S'INSCRIT L'ETUDE

Quand il s'agit d'évaluations par questionnaire en phase de « feed-back », les chefs de projet, commanditaires des évaluations, formulent généralement une demande peu spécifiée : évaluer les habitudes d'usage (ex. fréquence d'usage) et la perception d'une population d'utilisateurs donnée (ex. les jeunes de 18-25 ans) relatives à une famille de produits/services (ex. les sites Internet portail de France Télécom et leurs concurrents). Les résultats doivent

constituer des indicateurs qui leur permettent de prendre des décisions appropriées quant à la révision de leur produit/service (re-conception du produit, révision du service clientèle, etc.). Les dimensions à évaluer étant peu spécifiées, les professionnels de l'évaluation cherchent à produire une liste exhaustive de dimensions à partir de deux types de données : celles produites par des experts, enrichies de celles produites par des utilisateurs.

Le choix épistémologique est proche de celui observé dans l'évaluation sensorielle des produits alimentaires. Dans ce domaine, les mesures sensorielles sont réalisées par des personnes formées à l'évaluation (pour éviter les biais dus à la situation de laboratoire avec des sujets naïfs) et complétées par des mesures hédoniques simples, recueillies auprès de consommateurs naïfs (ou même des données sur la consommation). De même, pour l'évaluation des produits/services de France Télécom, les dimensions d'évaluation élaborées par des experts constituent une première base, à partir desquelles les données utilisateurs servent à ajouter ou à reformuler des dimensions.

Le processus de l'extraction des dimensions à la construction du questionnaire suit donc les quatre phases suivantes :

- 1. élaboration de dimensions experts : des dimensions sont d'abord élaborées par des experts de divers domaines du produit/service (évaluation, marketing, conception, etc.) à partir de leurs connaissances des utilisateurs/clients du produit/service. Cela se réalise sur la base d'un travail collectif entre eux ou à partir d'entretiens individuels menés et analysés par les professionnels de l'évaluation.
- 2. enrichissement des dimensions experts avec des données utilisateurs : des entretiens (le plus souvent individuels) sont menés avec différents groupes d'utilisateurs de la famille du produit/service à étudier (ex. les internautes de 18-25, 25-45, 45-55 et 55-70 ans) pour obtenir la plus large variété d'opinions possibles. Les données sont analysées de façon à déterminer ce qui peut enrichir la liste de dimensions experts. La nouvelle liste obtenue est donc une liste exhaustive.
- 3. réduction de la liste exhaustive : un large échantillon du seul groupe d'utilisateurs à étudier dans l'évaluation finale (ex. les internautes de 18-25 ans) est invité à hiérarchiser les dimensions selon leur degré d'importance.
- 4. construction du questionnaire d'évaluation : le questionnaire d'évaluation est construit à partir des seules dimensions les plus importantes.

Notre étude correspond à la seconde étape de ce processus. La première étape n'a pas eu lieu dans le cadre de cette expérimentation car la liste de dimensions experts, déjà réalisée,

nous est fournie par le service d'évaluation. Nous l'avons utilisée, de même que les professionnels de l'évaluation, comme point de départ.

L'exercice de hiérarchisation des dimensions par un échantillon de sujets spécifiques et la construction du questionnaire n'ont pas été réalisés non plus. En effet, cette étude ne s'inscrit pas dans une demande de production de résultats d'évaluation, initiée par des professionnels destinataires, mais dans une demande de production de connaissances par les professionnels de l'évaluation concernant les deux techniques utilisées dans la seconde étape de leur processus.

2.2. LES QUESTIONS RENSEIGNEES DANS L'ETUDE

Cette étude cherche à renseigner tant les interrogations des professionnels de l'évaluation que les questions de recherche évoquées au chapitre 2. Nous traitons d'abord les questions relatives aux contenus produits par chaque technique, puis celles relatives à leur productivité.

Dans le paragraphe 4, nous traitons les questions suivantes :

- les deux techniques apportent-elles des éléments nouveaux par rapport aux dimensions élaborées par les experts ? quelle technique génère le plus d'éléments nouveaux ?
- quelle technique produit le plus d'idées détaillées ?
- quelles sont la quantité et la variété des dimensions générées par chaque technique ?
- quelle technique produit le plus d'idées qui n'émergent pas dans l'autre technique ?
- quelle est la nature des points essentiels, partagés par plusieurs sujets participant dans chaque technique ?
- quelle technique produit le moins d'idées difficiles à interpréter ?

Concernant l'avant dernier point, chercher à le renseigner sert aussi à répondre à une question opérationnelle, spécifique à la situation des professionnels de l'évaluation. En effet, nous cherchons à savoir si les *focus groups* offrent la possibilité de réduire leur processus d'évaluation, notamment de supprimer la troisième étape, c'est-à-dire d'obtenir directement et rapidement une liste synthétique de dimensions sur laquelle baser la construction du questionnaire.

Après cela, dans le paragraphe 5, nous traitons deux questions relatives à la productivité des techniques :

- quelle technique est la plus productive en termes de ressources humaines mobilisées ou en termes de coûts financiers engagés pour leur réalisation ?
- à partir de quel nombre de sujets, les productivités en termes de ressources humaines et de coûts financiers déclinent-elle ?

Dans le paragraphe 3 ci-dessous, nous décrivons les sujets de l'étude et les méthodologies conçues pour chaque technique, puis dans chacun des paragraphes 4 et 5, nous présentons l'analyse des données et les résultats qui en sont issus.

3. DESCRIPTION DE L'ETUDE

De même que dans l'étude de Fern (1982), le principe de cette étude repose sur la comparaison des performances d'agrégats de huit entretiens individuels et de celles de *focus groups* de huit participants. Nous décidons d'utiliser sensiblement la même méthodologie ainsi que les mêmes populations de sujets dans l'une et l'autre technique, afin de se placer dans les conditions optimales de comparaison.

3.1. LES SUJETS

Partant du principe donné dans le processus élaboré par les professionnels de l'évaluation, nous avons choisi deux groupes de cyber-consommateurs afin d'obtenir une plus large variété d'opinions quant aux sites marchands :

- des utilisateurs « plus expérimentés » dans l'utilisation d'Internet et des sites marchands : deux à sept ans d'utilisation d'Internet et plusieurs achats avec paiement en ligne, dans les 12 derniers mois,
- des utilisateurs « plus novices » : un à deux ans d'utilisation d'Internet et moins de deux achats sans paiement en ligne, dans les 12 derniers mois.

Un agrégat de huit entretiens individuels avec des utilisateurs de la première population a été réalisé et un *focus group* avec huit autres participants de cette même population. De même, un agrégat de huit entretiens individuels avec des utilisateurs de la seconde population a été réalisé et un *focus group* avec huit autres participants de cette même population.

Pour chaque groupe de huit sujets, nous avons aussi choisi une répartition d'hommes et de femmes sur quatre tranches d'âge (18-25, 25-45, 45-55 et 55-70 ans) toujours pour une

meilleure représentativité de la population internaute grand public. Enfin, les sujets ont été choisis parmi une variété de catégories socioprofessionnelles, de façon à peu près équilibrée dans chaque groupe (personnes au foyer, étudiants, employés, professions intermédiaires, cadres et professions intellectuelles, retraités). Cependant, ils ont aussi tous été choisis à partir de leur facilité et volonté à s'exprimer sur le thème donné. Cette dimension est la plus importante quand il n'est pas possible de réaliser des groupes homogènes (Merton *et al.*, 1956), afin de minimiser les blocages qui peuvent se produire lors d'un *focus group* (ex. personnes qui restent silencieuses, voire conflits).

3.2. LA METHODOLOGIE

La méthodologie est basée sur le recueil de verbalisations, d'abord sans utilisation, puis avec utilisation de produits/services, afin d'obtenir des données verbales variées et riches.

3.2.1. Verbalisations avec et sans utilisation de produits/services

Une méthodologie élaborée lors de précédentes évaluations exploratoires (sur les outils de messagerie par Internet, sur les sites Internet portail) a servi de point de départ et a été modifiée sur la base d'observations empiriques. Nous y avons aussi incorporé des techniques d'élicitation décrites au chapitre 1.

Nous avons remarqué que les sujets étaient très absorbés par leur séquence d'utilisation lors de verbalisations consécutives à celle-ci. Très naturellement ils ne détachaient plus leur attention des détails relatifs aux éléments de l'interface qu'ils venaient d'employer (la couleur d'une page, l'icône d'un bouton d'action, etc.). Ces verbalisations trop ciblées sont problématiques dans le cadre de l'élicitation d'idées et pensées utiles à la construction d'un questionnaire en vue d'une étude exploratoire qui doit porter sur des aspects aussi bien ergonomiques qu'hédoniques et relatifs au marketing. Ainsi, nous avons étoffé la première partie de la précédente méthodologie sans utilisation de produits/services, en utilisant la technique d'élicitation associative et l'entretien d'explicitation.

La seconde phase, avec mise en situation d'utilisation de produits/services, a été maintenue car elle présente un intérêt considérable : elle permet d'analyser les verbalisations des sujets en référence à leur activité et, parfois, de préciser des éléments incertains, évoqués lors de la première phase (le vocabulaire de certains sujets est assez approximatif, en

particulier concernant la désignation d'éléments graphiques ou techniques). Ces deux phases, avec et sans utilisation, se complètent.

La seconde phase utilise la méthode des verbalisations libres issue de l'approche systémique de la psychologie appliquée à la qualité perçue (sans prendre en compte la méthode d'analyse des données qui sert à obtenir un portrait verbal d'un produit, puisque telle n'est pas notre finalité). Nous utilisons seulement la technique de production de verbalisations libres.

L'ensemble de cette méthodologie (chacune de ses phases) est adapté selon que la passation est individuelle ou collective.

3.2.2. Méthodologie de l'entretien individuel

La méthodologie de l'entretien individuel se déroule selon les cinq étapes suivantes⁸³, décrites ci-après :

- Phase de verbalisation sans utilisation de produits/services :
 - Etape 1 : la technique d'élicitation associative,
 - Etape 2 : l'entretien d'explicitation,
 - Etape 3 : une phase d'écriture de fiches-critère,
- Phase de verbalisation avec utilisation de produits/services :
 - Etape 4 : la méthode des verbalisations libres,
 - Etape 5 : une nouvelle phase d'écriture de fiches-critère.

3.2.2.1. Phase de verbalisation sans utilisation de produits/services

- Etape 1 : la technique d'élicitation associative

L'entretien individuel démarre par une consigne invitant le sujet à réagir à un stimulus verbal (les aspects positifs et négatifs des sites marchands) en écrivant chaque mot ou idée qui lui vient immédiatement à l'esprit. Après environ cinq minutes, l'intervieweur demande au sujet de lire la liste qu'il a rédigée et une courte discussion permet d'étoffer ou d'éclaircir ces premiers éléments.

- Etape 2 : l'entretien d'explicitation

Après cette première étape, un entretien d'explicitation est mené à partir d'une expérience précise. L'intervieweur demande au sujet de se souvenir de la dernière fois qu'il a

utilisé un site Internet marchand ou d'une autre expérience qu'il a envie d'évoquer et la déroule avec lui, pas à pas.

- Etape 3 : une phase d'écriture de fiches-critère

Une dernière phase, écrite, est destinée à faire faire la synthèse des idées formulées jusque là. L'intervieweur commence par présenter une analogie possible avec le téléphone mobile (ex. critères d'appréciation relatifs au coût, à l'esthétique, au poids, etc). Le sujet est ensuite invité à écrire les intitulés et éléments de définition (portant sur des éléments ou caractéristiques d'éléments constituant les sites marchands) qui correspondent à ses critères de jugement, sur des fiches vierges. Les sujets qui n'osent pas ou ont des difficultés à élaborer leurs idées par écrit doivent être davantage accompagnés.

L'exercice présente l'intérêt d'obliger le sujet à construire davantage sa pensée et à avoir une démarche plus analytique pour tenter d'élaborer chacun des critères qui sous-tendent sa perception et ses actions.

3.2.2.2. Phase de verbalisation avec utilisation de produits/services

- Etape 4 : la méthode des verbalisations libres

Le sujet est invité à utiliser deux sites marchands et à verbaliser librement ses impressions et les différences et similitudes qu'il observe.

Pour ce faire, il est installé face à deux ordinateurs et choisit un secteur de vente parmi une liste de quatre en tout (voyages, vêtements, produits culturels, vins) afin d'utiliser deux sites « pré-choisis » dans le secteur retenu. Le sujet est invité à réaliser une première tâche d'exploration de la page d'accueil de l'un, puis de l'autre site, à verbaliser tout haut et librement ses impressions, puis les différences et similitudes qui lui sont apparues. Il accomplit ensuite une tâche de recherche du produit de son choix en la réitérant (même produit à chaque fois) sur les deux sites. La tâche en question comprend également le remplissage des bons de commande mais s'arrête avant la fin de l'opération de commande, étant donné que nous ne disposons pas de moyens financiers pour offrir le produit. De même, il est invité à verbaliser tout haut ses impressions, puis les différences et similitudes qui lui sont apparues et enfin à préciser le site qu'il a préféré et les raisons de ce choix.

Au cours de l'accomplissement des diverses tâches, le sujet a été relancé de façon neutre aussi souvent que nécessaire⁸⁴ (ex. « Pouvez-vous dire tout haut vos impressions »). Nous

avons aussi encouragé le sujet à verbaliser tout haut ce qu'il fait pour mieux comprendre la logique de ses agissements seulement quand elle nous est vraiment obscure. Enfin, nous nous sommes assuré qu'il compare des pages de même niveau sur les deux sites car parfois il compare, par exemple, la fiche de description détaillée du produit et la liste avec plusieurs produits peu détaillés, issue de la recherche par mots clefs.

Au total, huit sites marchands sont parcourus de façon aléatoire⁸⁵ selon les choix des sujets.

- Etape 5 : une nouvelle phase d'écriture de fiches-critère

Cette phase de verbalisations libres se clôture par une dernière phase d'écriture de fiches-critère. Une autre série de fiches vierges est présentée au sujet, sur lesquelles il est libre de proposer de nouveaux critères ou de reformuler les mêmes que précédemment en y ajoutant de nouveaux éléments de définition.

L'ensemble de l'entretien individuel dure une heure et demie.

3.2.3. Méthodologie du focus group

Comme nous l'avons vu au chapitre 1, le *focus group* est, à l'origine, une discussion modérée par un animateur qui cherche à recadrer les interactions libres quand nécessaire. Nous avons aussi vu que le *focus group* est propice à l'intégration de techniques collectives d'analyse ou de créativité. Ainsi, la technique d'élicitation associative, l'entretien à partir d'une expérience vécue et la méthode des verbalisations libres sont utilisées et adaptées à la situation de groupe. La méthodologie de l'entretien individuel est conservée dans la mesure du possible, c'est-à-dire compte tenu des contraintes que présente la situation de groupe.

La méthodologie du *focus group* se déroule selon les cinq étapes suivantes⁸⁶, décrites ci-après :

- Phase de verbalisation sans utilisation de produits/services :
 - Etape 1 (et 3⁸⁷): la technique d'élicitation associative et la création et nomination de critères collectifs,
 - Etape 2 : discussions de groupe à partir d'expériences vécues, équivalant à l'entretien d'explicitation,
 - Etape 3 : cf. étape 1 du *focus group* (création et nomination de critères collectifs) ,
- Phase de verbalisation avec utilisation de produits/services :

- Etape 4 : la méthode des verbalisations libres, adaptée à la situation de groupe,
- Etape 5 : discussion de synthèse et établissement d'une seconde liste de critères collectifs.

3.2.3.1. Phase de verbalisation sans utilisation de produits/services

- Etape 1 (et 3) : la technique d'élicitation associative et la création et nomination de critères collectifs

Le *focus group* démarre par une phase individuelle d'écriture rapide sur les aspects positifs et négatifs des sites marchands. Après environ cinq minutes, chaque sujet, un par un, énonce ses différents éléments (appelés éléments TEA) à tout le groupe et l'animateur les note sur un « paper-board » en les numérotant. Krueger et Casey (2000) préconisent aussi de démarrer un *focus group* par une séquence de travail individuel pour avoir une trace des idées propres à chaque participant et pour aider à identifier les éventuels biais relatifs aux effets de leadership lors des interactions ultérieures de groupe.

Cette première phase courte s'apparente aussi à la technique du *nominal group* (écriture ou réponse à des questions de façon individuelle ; Stewart & Shamdasani, 1990), mais au lieu de la phase de vote qui s'en suit généralement (Langford & McDonagh, 2003b), un travail collectif de catégorisation a été réalisé. Les éléments TEA sont pris, un par un, dans l'ordre et regroupés par similitude pour élaborer des critères de jugement auxquels les participants donnent un titre.

Parmi environ une trentaine d'éléments TEA écrits sur le « paper-board », les cinq premiers sont d'abord retenus. Leur numéro est inscrit sur une fiche cartonnée et ces fiches placées, de façon éparse, sur le tableau « autocollant » qui va constituer « l'espace de travail » du groupe (Urdapilleta, 2001b). L'animateur invite alors les participants à regrouper les éléments TEA similaires, s'il y en a. Sinon, les cinq éléments TEA suivants sont considérés, leur numéro inscrit sur des fiches et le même travail de regroupement par similitude accompli.

Des discussions ont lieu sur les similitudes et différences de chaque élément par rapport aux autres, lors desquelles l'animateur insiste sur les points de convergence et reformule pour les clarifier les points de divergence. Les participants sont amenés à trouver un consensus autour du regroupement de chaque élément TEA avec un ou plusieurs autres. Si un consensus n'est pas trouvé, l'élément est laissé de côté et reconsidéré ultérieurement. Enfin, des regroupements sont parfois modifiés au fur et à mesure que de nouveaux éléments TEA sont introduits. Ce travail est similaire à celui de la réalisation de *word map* (cf. chapitre 1) ou

aussi d'*affinity diagram* par des professionnels du marketing ou de la conception pour travailler à partir d'un ensemble de données utilisateurs/clients (ex. Holtzblatt & Beyer, 1993; Urban & Hauser, 1993).

Les participants sont aussi invités à donner un titre aux regroupements. Une fois les regroupements et leurs titres stabilisés, ces derniers constituent l'équivalent des précédentes fiches-critère (cf. étape 3 de la méthodologie de l'entretien individuel).

- Etape 2 : discussions de groupe à partir d'expériences vécues, équivalant à l'entretien d'explicitation.

Le précédent travail de catégorisation suscite des discussions où les participants exposent leurs différents points de vue, très souvent à l'appui d'expériences vécues qu'ils explicitent. Cela remplace l'entretien d'explicitation puisqu'il est impossible de mener un tel entretien avec chacun des participants⁸⁸.

Ces échanges génèrent aussi spontanément des recherches de consensus dans le groupe ou au contraire l'établissement de désaccords entre les participants concernant les différences ou similitudes entre éléments TEA.

- Etape 3 : cf. étape 1 de la méthodologie du *focus group*.

3.2.3.2. Phase de verbalisation avec utilisation de produits/services

- Etape 4 : la méthode des verbalisations libres, adaptée à la situation de groupe.

La méthode des verbalisations libres est habituellement réalisée en passation individuelle, cependant ici elle est adaptée à la situation de groupe. Les participants choisissent un des quatre secteurs de vente possibles, et un volontaire navigue selon les décisions du groupe (depuis un seul micro-ordinateur portable), en réalisant la tâche d'exploration de la page d'accueil de l'un, puis de l'autre site marchand, puis la tâche de la recherche d'un produit également sur l'un et l'autre site. L'image du site marchand est projetée sur un grand écran. Les autres participants indiquent au surfeur que parcourir sur le site et choisissent le produit à chercher (de courtes discussions sont consacrées à la prise de décisions collectives). Etant donné qu'un seul participant navigue sur le site, cela fonctionne davantage comme une démonstration visuelle du produit (Langford & McDonagh, 2003b). Par ailleurs, suite à la tâche de recherche, le formulaire des bons de commandes ne sont pas remplis, pour plus de simplicité et pour une question de temps.

Dès les pages Internet à l'écran, de nombreux commentaires sont énoncés, des expériences personnelles partagées et discutées, etc. Les échanges donnant lieu à des discussions d'aparté ou étant parfois hors sujet doivent être régulièrement recadrés par l'animateur. Les participants sont aussi invités à commenter les différences et similitudes observées, suite à chacune des tâches accomplies sur les deux sites.

- Etape 5 : discussion de synthèse et établissement d'une seconde liste de critères collectifs.

Une rapide discussion de synthèse leur est ensuite proposée pour faire le point sur les critères de jugement relatifs à cette séquence d'utilisation/observation. L'animateur recherche les consensus dans le groupe et dresse une liste de critères sur le « paper-board », ce qui fait office de seconde liste de fiches-critère (cf. étape 5 de la méthodologie de l'entretien individuel).

L'ensemble du *focus group* dure deux heures et demie.

3.3. L'ANIMATEUR OU INTERVIEWEUR

La même personne a animé les *focus groups* et interviewe les sujets en face-à-face, toujours pour des raisons de comparaison optimales. Ce professionnel présente une expérience conséquente d'animation d'entretiens collectifs (plusieurs années) et est formé à l'entretien d'explicitation. Les deux autres techniques lui ont été décrites en détail par écrit et lors d'échanges de mise au point de l'étude.

Deux entretiens individuels ont été réalisés en pré-test afin que l'intervieweur s'aguerrisse à l'ensemble de la méthodologie. Concernant les *focus groups*, la méthodologie a été peaufinée grâce aux échanges avec ce professionnel et grâce aux trouvailles du technicien (immense tableau « autocollant » pour placer et déplacer les fiches cartonnées).

3.4. LE DISPOSITIF DES ENTRETIENS

Les entretiens individuels et collectifs se sont tenus dans deux salles, respectivement, l'une petite, l'autre grande, d'un laboratoire d'évaluation indépendant des locaux de France Télécom pour ne pas influencer les participants. Des micro-ordinateurs avec une connexion Internet (et les sites Internet marchands nécessaires en favoris dans le navigateur) ont été installés dans chacune de ces salles, aménagées de façon à reproduire l'ambiance la plus conviviale et familière possible. Deux caméras et des microphones, gérés depuis la salle de régie, ont permis de recueillir les enregistrements sonores et audio-visuels de chaque entretien.

Dans la grande salle, pour des raisons d'espace, plusieurs rangées de sièges accueillent les participants même si une table ronde est généralement plus indiquée. En effet, il faut que

tous les participants puissent regarder (sans avoir à se tourner) le « paper-board » et le tableau « autocollant », puis l'écran sur lequel sont projetés les deux sites marchands.

Enfin, un dispositif constitué de deux ordinateurs en réseau permet au chercheur dirigeant l'étude, d'échanger avec l'intervieweur depuis la salle de régie. Cela, pour des raisons évidentes de double-tâche dans l'activité de l'intervieweur, ne s'est réalisé que très rarement et seulement lors des entretiens individuels, quand des oublis évidents se sont produits lors des phases d'écriture des fiches-critère.

Après avoir exposé la méthodologie employée dans chaque technique, nous présentons l'analyse des données et les résultats, concernant d'abord les contenus produits, puis concernant la productivité de chaque technique.

4. DONNEES ET RESULTATS SUR LA PRODUCTION DE CONTENU VERBAL ET DE DIMENSIONS

4.1. ANALYSE DES DONNEES

Nous présentons d'abord la constitution de la grille d'analyse des verbalisations qui correspond à la liste exhaustive de dimensions utilisateurs, puis le codage et d'autres éléments d'analyse des données.

4.1.1. Elaboration d'une liste exhaustive de dimensions utilisateurs

Nous avons élaboré une grille globale d'analyse de toutes les verbalisations, c'est-à-dire celles produites par les deux agrégats d'entretiens individuels aussi bien que celles produites par les deux *focus groups*. Cela produit une liste exhaustive de dimensions utilisateurs qui concernent des aspects ergonomiques, mais aussi esthétiques, commerciaux, etc.

Nous avons utilisé une liste de dimensions experts comme trame de départ pour traiter ces données. Cette liste de référence nous est fournie par le service d'évaluation. Elle a été élaborée par trois professionnels de ce service, de formation ergonomique ou linguistique, et par un ergonome impliqué dans la conception de sites marchands. Ces professionnels ont travaillé collectivement de façon à synthétiser leurs connaissances des internautes et cyber-consommateurs.

Les entretiens utilisateurs ont été retranscrits et les enregistrements audio-visuels visionnés pour analyser les verbalisations des utilisateurs corrélativement à leur activité concernant la seconde partie de chaque entretien. Une analyse de contenu par thème a été réalisée, en partie, en référence à l'activité des utilisateurs. Des catégories d'analyse correspondant à l'utilisation et perception associée des différents objets (perçus ou verbalisés) constituant un site marchand ont été déterminées. Cependant, une taxonomie exhaustive des opérations réalisables (objets et actions) dans le contexte des tâches possibles sur un site marchand est impossible à définir par avance (Bisseret *et al.*, 1999). Cette taxonomie nous est en partie donnée par la liste des 24 dimensions experts (cf. tableau 2 ci-après). Les données de tous les entretiens utilisateurs nous ont ensuite amené à réorganiser en partie cette liste, puis à l'enrichir. Certaines dimensions d'origine ont été scindées en plusieurs dimensions. Finalement, 31 catégories d'analyse ou dimensions ont été dégagées. Elles constituent la liste exhaustive de dimensions utilisateurs (cf. étape 2 du processus d'évaluation en début de chapitre). Enfin, la littérature sur les sites marchands nous a aidé à consolider cette grille d'analyse.

Tableau 2 : Liste de dimensions élaborées par les experts concernant la perception des sites Internet marchands par les internautes et cyber-consommateurs

N° de dimension	Intitulé de la dimension
1	Facilité d'accès au site
2	Renommée du site
3	Positionnement marketing du site
4	Compréhension immédiate du service rendu par le site dès la page d'accueil
5	Qualité et intérêt des produits proposés
6	Quantité des produits proposés
7	Prix des produits proposés
8	Utilité des services associés à l'achat de produits (livraison, emballage cadeau...)
9	Coût des services associés
10	Qualité des informations sur les produits
11	Clarté des informations sur les conditions de paiement et modalités de livraison
12	Sécurisation de la transaction
13	Facilité d'utilisation des procédures de commande
14	Personnalisation du contenu
15	Suivi de la commande
16	Une relation humaine au-delà de la machine
17	Accessibilité et efficacité de la chaîne de distribution
18	Compatibilité avec les logiciels courants (navigateurs...)
19	Simplicité de l'organisation de l'information
20	Qualité de la conception du site (visibilité des liens, des commandes et accès principaux ; existence d'un formulaire de recherche multi-choix ; rapidité de chargement ; liens d'accès à des informations complétant chaque produit)
21	Clarté du vocabulaire
22	Gestion des erreurs
23	Esthétique du site
24	Pas d'excès dans la publicité en ligne

Les 31 nouvelles dimensions (voir l'extrait de liste exhaustive de dimensions utilisateurs dans le tableau 3 ci-après) correspondent à un objet générique matériel ou symbolique (ex. notoriété, service rendu, informations sur les produits, etc.) et, parfois, à une ou plusieurs caractéristiques associées (clarté, qualité, facilité). Ces caractéristiques sont parfois sous-entendues. Par exemple, concernant la catégorie 1, il s'agit bien sûr d'une bonne notoriété car les internautes sont alors rassurés, notamment quant à la qualité des produits vendus. Les actions ne sont pas précisées à ce niveau car plusieurs sont possibles avec chaque objet (ces catégories sont génériques).

Tableau 3 : Extrait de la liste exhaustive de dimensions utilisateurs, équivalant à la grille d'analyse des verbalisations des entretiens individuels et des focus groups sur la perception des sites Internet marchands⁸⁹

N° de dimension	Intitulé de la dimension
1	Notoriété
2	Facilité d'accès
3	Clarté du service rendu et informations qui rassurent dès la page d'accueil
4	Accès et qualité du contenu des informations de vente et de sécurité (avant la commande)
5	Existence et qualité de la relation humaine au-delà de la machine (conseiller en vente)
6	Diversité et qualité des produits en vente
7	Prix des produits proposés
10	Qualité de la description et des informations relatives aux produits (caractéristiques, photo, disponibilité, prix, etc.)
11	Existence de différentes fonctions de recherche (thématique de départ)
12	Facilité de la recherche de produits (saisie et cheminement)
13	Pertinence et facilité de lecture des résultats de la recherche
14	Informations supplémentaires et propositions d'autres produits
15	Facilité de la sélection des produits à la commande (panier)
16	Facilité et rapidité de la commande (inscription/commande/paiement)

Les éléments écrits par les utilisateurs dans les fiches-critère et des indications concernant leur activité d'utilisation ont été ajoutés aux segments verbaux correspondant, issus de l'analyse de contenu. Ces ensembles « verbalisations - éléments écrits – indication sur l'activité » constituent les unités sémantiques codées. Les 16 entretiens individuels et les deux *focus groups* ont produit en tout 470 unités sémantiques. Le tableau 4 ci-dessous présente trois exemples d'unités sémantiques.

Tableau 4 : Trois unités sémantiques obtenues après analyse de contenu des entretiens individuels et des focus groups sur la perception des sites Internet marchands

Unité sémantique 1. Possibilité de différentes clefs de recherche : « <i>Pour les spectacles : un tri par genre, un tri géographique, un tri alphabétique... le choix multiple est très aidant. Les sous-menus déroulants sont en général pratiques plus que d'exposer, par exemple, page après page une série de produits</i> ».
Unité sémantique 2. Sous forme de livre avec onglet – par rubrique : « <i>C'est fait sous forme de livre dont on peut tourner les pages, en cliquant avec un petit onglet ; c'est agréable</i> ».
Unité sémantique 3. Seulement 2 livres ! (utilisation de sites de produits culturels) : « <i>Il n'y a que deux livres sur la seconde République. Cela me paraît léger, je pensais qu'il y en aurait plus que ça</i> ».

Légende : les unités sémantiques sont constituées d'éléments écrits par les utilisateurs (en style normal), de leurs verbalisations (en italique) et éventuellement d'indications sur leur activité (entre parenthèses).

Les unités sémantiques ont donc été regroupées au fur et à mesure en catégories, c'est-à-dire dimensions qui elles-mêmes ont pu être découpées en sous-catégories, les sous-dimensions, elles-mêmes découpées en sous-catégories, les détails. En effet, en fonction du niveau de précision de l'information contenue, les 470 unités sémantiques ont aussi pu être séparées en trois catégories : celles du niveau de la dimension, celles du niveau de la sous-dimension, celles du niveau du détail. Cela correspond à un second critère de segmentation (Bisseret *et al.*, 1999).

La grille d'analyse des verbalisations ou liste exhaustive des dimensions utilisateurs est, en fait, constituée de :

- 31 dimensions correspondant à une information de niveau général, c'est-à-dire qui concerne un objet générique et une ou plusieurs caractéristiques de cet objet (ex. diversité et qualité des produits en vente ; existence de différents moyens de recherche⁹⁰),
- 110 sous-dimensions correspondant à une information plus précise, c'est-à-dire qui concerne une caractéristique précise d'une classe d'objets (ex. diversité des produits en vente, originalité des produits en vente, qualité...⁹¹). A ce niveau des actions peuvent aussi être précisées (ex. possibilité de faire une recherche par rubriques et sous-rubriques⁹²),
- 301 détails correspondant à une information d'un niveau très précis, c'est-à-dire qui concerne une caractéristique ou l'existence d'un objet précis (ex. appréciation de plusieurs univers/départements de produits sur un site; appréciation de la diversité des produits dans une même gamme - ex. plusieurs marques dans le rayon appareils photos⁹³). A ce niveau, des actions plus précises peuvent aussi être exprimées (possibilité de faire une recherche par rubriques pour la possibilité d'un affinement successif⁹⁴).

Le tableau 5 ci-après montre un extrait de la liste de dimensions utilisateurs, de leurs sous-dimensions et de leurs détails.

Tableau 5 : Extrait de la liste exhaustive détaillée de dimensions utilisateurs, équivalant à la grille détaillée d'analyse des verbalisations des entretiens individuels et des focus groups

Dimension	Sous-dimensions	Détails
6. Diversité et qualité des produits en vente	6a. Diversité des produits/services et gamme de produits - Unité sémantique 3	6a1. Appréciation de plusieurs univers/départements de produits/services sur un site (ex. musique, livres, hifi, etc.)
		6a2. Appréciation de la diversité des produits/services dans une même gamme (ex. plusieurs marques au rayon appareils photos), ou de la diversité des tailles, coloris, éditeurs, etc.
	6b. Originalité des produits/services	6b1. Appréciation de produits/services en exclusivité sur Internet (ex. petits déjeuners à domicile)
		6b2. Appréciation de produits/services spécifiques, rares (ex. produits étrangers)
	6c. Qualité des produits/ services	6c1. Appréciation d'une bonne qualité des produits/services (aussi en rapport avec l'origine ou la marque des produits).
11. Existence de différents moyens de recherche	11a. Possibilité de recherche par rubriques (recherche hypertexte, de lien en lien)	11a1. Appréciation du système d'onglets principaux - Unité sémantique 2
		11a2. Appréciation de la recherche par rubriques pour la possibilité d'affinement successif
		11a3. Appréciation de la deuxième barre de navigation (après avoir choisi un onglet principal) avec des récapitulatifs permanents de sous-rubriques
	11b. Possibilité de recherche par mots clefs	11b1. Appréciation de l'obtention directe des résultats par la simple saisie de mots-clefs
		11b2. Appréciation de l'existence de l'association du système d'onglets + saisie de mots-clefs (sur chaque page Internet).
		11b3. Appréciation de la double-entrée dans le champ de saisie (mots clefs et présélections dans des sous-menus déroulants)
	11c. Possibilité de recherche par formulaire multi-choix	11c1. Appréciation de l'association de plusieurs critères de choix (surtout utile pour les voyages, les spectacles)- Unité sémantique 1
		11c2. Appréciation de la possibilité de recherche par le critère du prix (ex. sites de vins)
	11d. ...	11d1., 11d2., ...

La suite de l'analyse des données, en particulier le codage des unités sémantiques, est basée sur les critères pour la comparaison des deux techniques et qui sont relatifs aux questions à renseigner (cf. paragraphe 2.2. au début du chapitre).

Pour des raisons de compréhension de l'analyse des données, elle sera décrite dans l'ordre présenté ci-dessous :

- Critères qui concernent la nature des données en entretien individuel versus en *focus group* :
 - 1. Variété de dimensions différentes et quantité d'idées (différentes) détaillées,
 - 2. Quantité d'idées nouvelles par rapport aux dimensions élaborées par les experts,
 - 3. Quantité d'idées spécifiques à chaque technique,
 - 4. Points essentiels dégagés dans chaque technique (dimensions collectives ou individuelles partagées),
- Critère qui concerne l'exploitabilité des données en entretien individuel versus en *focus group* :
 - 5. Quantité d'idées difficiles à interpréter.
- Comparaison de la situation de verbalisation avec ou sans utilisation de sites marchands :
 - 6. Quantité d'idées nouvelles, détaillées, spécifiques et difficiles dans chaque situation de verbalisation.

Enfin, nous comparons les deux agrégats d'entretiens individuels et les deux *focus groups* sans tenir compte de la variable « population », c'est-à-dire de la différence relative aux sujets « plus novices » ou « plus expérimentés ». Dans chacune des deux techniques, la même proportion de participants est issue de l'une et de l'autre population d'internautes. L'effet de cette variable est neutralisé⁹⁵.

4.1.2. Description de l'analyse des données et du codage des unités sémantiques

4.1.2.1. La variété de dimensions différentes et les idées (différentes) détaillées

Naturellement, il est arrivé que plusieurs sujets énoncent la même idée. Afin d'identifier la variété des idées, notre objectif est de ne retenir que les idées ou unités sémantiques différentes et non les occurrences de chacune. Dans chaque agrégat de huit entretiens individuels et dans chaque *focus group*, les idées identiques d'un individu à un autre n'ont donc été retenues qu'une seule fois, à un même niveau de profondeur. Par exemple, si une

idée, énoncée successivement par deux individus, correspondait à la sous-dimension « diversité des produits en vente », une seule unité sémantique de l'agrégat d'entretiens ou du *focus group* a été conservée (la première dans le protocole verbal). De même, si une idée énoncée par plusieurs des huit sujets/participants correspondait au même détail « appréciation de plusieurs univers/départements de produits sur un site ».

Cela a permis d'obtenir les idées qui sont toutes différentes sur le niveau des détails (idées détaillées), sur le niveau de la sous-dimension (idées non détaillées) et sur le niveau de la dimension (idées non détaillées aussi). Par exemple, les deux premières unités sémantiques du tableau 4 font partie de la dimension « Existence de différents moyens de recherche » (dimension 11, cf. tableau 5) et sont du niveau du détail sur cette dimension (11c1 et 11a1 respectivement, cf. tableau 5). Par ailleurs, la troisième unité sémantique fait partie de la dimension « Diversité et qualité des produits en vente » (dimensions 6) et est du niveau de la sous-dimension (6a). Finalement, la quantité d'idées détaillées versus non détaillées dans chaque technique a donc pu être mesurée.

Enfin, pour connaître la variété des aspects différents abordés dans chaque technique, les nombres entiers de dimensions (parmi 31 en tout) et de sous-dimensions (parmi 110 en tout) révélées dans chaque technique ont été établis. Dès qu'une unité sémantique relative, par exemple, au détail 11a1 a été exprimée, celle-ci faisant partie de la sous-dimension 11a, qui elle-même fait partie de la dimension 11, les dimension et sous-dimension 11 et 11a ont été révélées et comptabilisées pour l'agrégat d'entretiens ou le *focus group*. Néanmoins cette unité sémantique appartient bien au niveau du détail. Autrement dit, des dimensions peuvent être révélées mais les unités sémantiques participant à cela sont du niveau du détail ou de la sous-dimension.

4.1.2.2. Les idées nouvelles par rapport aux dimensions élaborées par les experts

Pour mesurer la quantité d'idées nouvelles (versus non nouvelles) dans chaque technique, nous avons spécifié si chacune des unités sémantiques était nouvelle ou non par rapport aux éléments de la liste de dimensions experts. Dans les exemples des unités sémantiques du tableau 4 (cf. aussi tableau 6 ci-après), seulement la seconde est nouvelle car elle est relative à la recherche par rubriques. Tout ce qui concerne la recherche hypertextuelle⁹⁶ et la recherche par mots-clefs⁹⁷ constitue des éléments qui n'existent pas dans les dimensions experts (cf. tableau 2). En revanche, les unités une et trois ne sont pas nouvelles. L'unité 1 est relative à la

recherche par formulaire multi-choix⁹⁸ qui est un élément de la dimension 20 de la liste des dimensions experts. L'unité 3 concerne la diversité des produits en vente⁹⁹ qui est un élément de la dimension 6 de la liste des dimensions experts.

De plus, les idées nouvelles qui étaient aussi des idées de détail ont été identifiées. Le tableau 6 ci-dessous montre le résumé du codage des trois précédents exemples d'unités sémantiques (les entretiens individuels sont mentionnés par EI et les *focus groups* par FG).

Tableau 6 : Résumé du codage de trois exemples d'unités sémantiques obtenues suite à l'analyse de contenu des entretiens individuels et des focus groups

	Idées nouvelles vs. non nouvelles	Idées détaillées vs. non détaillées	Idées spécifiques vs. non spécifiques (ou communes)	Idées difficiles vs. non difficiles	N° de codage correspondant au trois niveaux de profondeur
Unité 1	Ñ	D	S aux EI	Đif	11c1
Unité 2	N	D	S aux EI	Đif	11a1
Unité 3	Ñ	Đ	Š ou communes aux EI et aux FG	Đif	6a

Légende : N: idées nouvelles; Ñ: idées non nouvelles; D: idées détaillées; Đ: idées non détaillées; S: idées spécifiques; Š: idées non spécifiques; Dif: idées difficiles à interpréter; Đif: idées non difficiles.

4.1.2.3. Les idées spécifiques à chaque technique

Par ailleurs, pour savoir quelle technique apporte le plus d'idées qui n'émergent pas dans l'autre technique, la quantité d'idées spécifiques à chaque technique a été mesurée. Nous avons donc défini si chaque unité sémantique était spécifique à chaque technique ou communes aux deux (c'est-à-dire existante dans au moins une des deux passations de chaque situation). Par exemple, les deux premières unités sémantiques du tableau 4 (cf. aussi tableau 6) sont spécifiques aux entretiens individuels, elles n'ont pas été exprimées lors des *focus groups*. La troisième, au contraire, a été évoquée dans les deux techniques. Elle est issue des entretiens individuels et une autre unité sémantique d'un des *focus groups* exprime exactement la même idée.

4.1.2.4. Les méga-dimensions ou points essentiels générés par chaque technique

Nous avons aussi dégagé les points essentiels générés par chaque technique. Dans le cas des *focus groups*, cela correspondait directement aux titres des regroupements réalisés collectivement. Dans le cas des entretiens individuels, il a fallu construire ce qui était collectif

à partir des données individuelles, c'est-à-dire identifier les dimensions utilisateurs les plus partagées par les sujets de chaque agrégat d'entretiens.

A la suite de chaque *focus group*, les titres des regroupements ont été recueillis. Le recoupement des titres similaires entre le premier et le second *focus group* a permis d'en extraire une synthèse.

D'autre part, les dimensions utilisateurs (correspondant à celles de la grille d'analyse des verbalisations) les plus partagées par les huit interviewés en face-à-face ont été déterminées, après un codage systématique des verbalisations. La règle était la suivante : si quatre à huit sujets ont produit l'ensemble des unités sémantiques compris dans une dimension, la dimension a été retenue. En revanche, si la dimension a été évoquée par seulement trois sujets, elle a été écartée.

Etant donné leur caractère collectif ou partagé, ces dimensions ne correspondent pas aux dimensions utilisateurs de la liste exhaustive (grille d'analyse). Ce sont des méga-dimensions dans les deux cas.

4.1.2.5. Les idées difficiles à interpréter

Pour connaître la meilleure technique en termes d'exploitabilité de ces données, nous avons calculé la quantité d'idées difficiles à interpréter dans chacune. Nous avons donc établi si chacune des unités sémantiques faisait partie de la catégorie facile ou difficile à interpréter. Les données difficiles contiennent des éléments qui sont peu explicites et peuvent être compris seulement grâce à une bonne connaissance du domaine.

Les trois précédents exemples d'unités sémantiques ne sont pas des idées difficiles à interpréter (cf. tableau 6). Pour donner une illustration d'une unité difficile, voici un exemple (en style normal, ce qui est écrit, et en italique, ce qui est verbalisé) : « Raccourci - pour accéder à ce qu'on cherche : *Un lien vous voulez dire ? Oui, oui...* ». Il n'est pas très clair quel type de lien ou de raccourci est souhaité. Peut-être est-ce un souhait de davantage de liens sur le site (problème d'architecture du site) ou un souhait de liens de rappel des pages visitées à partir de la page en cours (et permettant de retourner sur ces pages, sans utiliser la commande du navigateur de retour à la page précédente) ?

4.1.2.6. La situation de verbalisation sans versus avec utilisation du site marchand

Enfin, il s'agit de comparer la situation de verbalisation sans utilisation du site marchand versus avec utilisation dans chaque technique. La situation de verbalisation sans utilisation correspond à la première partie de chaque entretien comprenant la technique d'élicitation associative, l'entretien à partir d'expériences vécues (et la phase d'écriture de synthèse) tandis que la situation de verbalisation avec utilisation correspond à la seconde partie avec la méthode des verbalisations libres (et la seconde phase d'écriture de synthèse). L'ensemble des précédents critères est repris pour cette comparaison. Les unités sémantiques, déjà codées, ont simplement été départagées selon qu'elles proviennent de la situation sans ou avec utilisation et les quantités d'idées nouvelles, détaillées, spécifiques et difficiles établies dans chacune de ces situations, dans chaque technique.

4.2. RESULTATS

Les résultats sont présentés selon l'ordre suivant :

- 1. Quantité d'idées nouvelles par rapport aux dimensions élaborées par les experts,
- 2. Variété de dimensions différentes et quantité d'idées (différentes) détaillées,
- 3. Quantité d'idées spécifiques à chaque technique,
- 4. Dimensions collectives versus individuelles partagées,
- 5. Quantité d'idées difficiles à interpréter,
- 6. Quantités d'idées nouvelles, détaillées, spécifiques et difficiles à interpréter dans la situation de verbalisation avec ou sans utilisation du site marchand.

Pour rappel, les entretiens individuels sont mentionnés par EI et les *focus groups* par FG dans les tableaux et figures ci-dessous.

4.2.1. Quantité d'idées nouvelles par rapport aux dimensions élaborées par les experts

Nous présentons tout d'abord le résultat concernant les idées nouvelles par rapport à la liste de dimensions experts. Le pourcentage d'unités sémantiques nouvelles (versus non nouvelles) par rapport au volume global d'unités de chaque technique (329 unités dans les deux agrégats d'entretiens individuels ou 171 dans les deux *focus groups*) permet de statuer sur l'intérêt des données utilisateurs (cf. tableau 8 au § 4.2.2.). Environ 52% de toutes les idées produites en entretien individuel relèvent d'informations que les experts n'ont pas

produites lors de leur travail collectif d'élaboration des dimensions. Par ailleurs, environ 37% de toutes les idées produites en *focus group* sont nouvelles par rapport aux dimensions experts. Cette différence est significative ($\chi^2=9,38$ et $p=0.022$).

Les entretiens individuels apportent donc significativement plus de données nouvelles que ne le font les *focus groups*. Toutefois, les deux techniques produisent une quantité de données nouvelles non négligeable. L'intérêt des données utilisateurs est établi et nous allons comparer les deux techniques plus finement.

4.2.2. Variété des dimensions et quantité d'idées détaillées

Le nombre entier de dimensions, de sous-dimensions et de détails révélés pour chaque passation (cf. tableau 7 ci-après) met au jour la variété des aspects abordés par chaque technique. Contrairement aux diverses quantités d'unités sémantiques (nouvelles, détaillées, etc.) qui peuvent être rapportées en pourcentage, il ne peut être révélé qu'un nombre entier de dimensions¹⁰⁰ de chaque niveau et cela pour chaque passation. (Il faut ainsi tenir compte des durées de chacun – 12 heures avec huit entretiens individuels versus deux heures et demie de *focus group*).

Tableau 7 : Comparaison de l'entretien individuel et du focus group en termes de dimensions révélées de trois différents niveaux de détail (dimensions, sous-dimensions et détails)

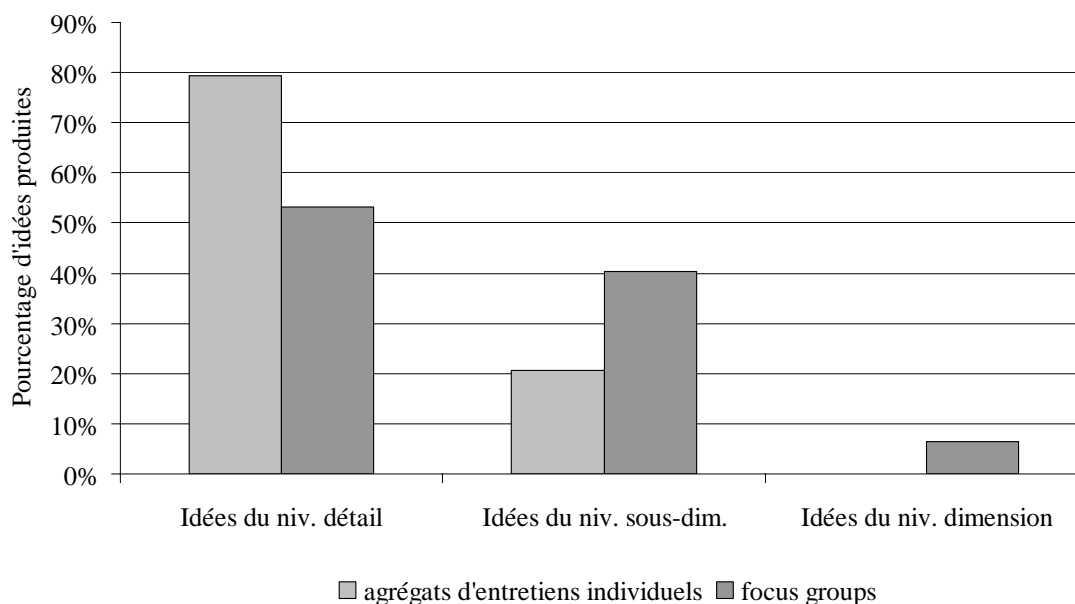
Sujets	Technique	Nombre d'entités révélées		
		Parmi 31 dimensions	Parmi 110 sous-dimensions	Parmi 301 détails
« Plus expérimentés »	EI	29	86	142
« Plus novices»	EI	30	80	119
« Plus expérimentés »	FG	24	50	43
« Plus novices»	FG	22	45	32

Même si, tout au plus, 30 dimensions sont révélées à la fois avec l'agrégat d'entretiens individuels des sujets « plus novices », le recouvrement entre les quatre passations fait que les 31 dimensions sont révélées. Aucune dimension de la liste élaborée par les experts ne demeure inexprimée dans les verbalisations des utilisateurs.

Un test t de Student à partir des moyennes de ces nombres entiers en entretien individuel, d'une part, et des moyennes de ces nombres entiers en *focus group*, d'autre part, montre que les différences à chaque niveau de profondeur sont significatives avec une marge d'erreur inférieur à 3% (respectivement, $t=5.81$, $\rho=0.028$; $t=9.09$, $\rho=0.011$; $t=7.29$, $\rho=0.018$). Ainsi, le nombre entier de détails en entretien individuel est beaucoup plus élevé qu'en *focus group*. Ce n'est pas le cas du nombre de sous-dimensions et encore moins de celui des dimensions, surtout si l'on prend en compte les durées d'entretien (24 heures d'entretien individuel en tout versus cinq heures de *focus group*). Les entretiens individuels semblent l'emporter concernant la production de détails. Les résultats présentés ci-après vont permettre de confirmer ou d'infirmier cela.

Le graphique de la figure 3 ci-après révèle la répartition en pourcentage (par rapport au volume d'unités de chaque technique) des unités du niveau de la dimension, de celles du niveau de la sous-dimension et de celles du niveau du détail.

Figure 3 : Comparaison de l'entretien individuel et du focus group en termes d'idées produites de trois différents niveaux de détail (dimensions, sous-dimensions, détails)



Légende : Il s'agit du pourcentage d'idées produites à chacun des niveaux de profondeur : niveaux de la dimension, de la sous-dimension, du détail.

Un test t de Student montre que les différences concernant les niveaux du détail et de la dimension seulement sont significatives (respectivement, $t=9.00$, $\rho=0.012$; $t=7.29$, $\rho=0.018$), mais pas celle du niveau de la sous-dimension ($t=1.64$, $\rho=0.24$). La figure 3 confirme donc que les entretiens individuels ont l'avantage sur les *focus groups* concernant les unités sémantiques de détail. Les idées détaillées représentent 79% de toutes les idées produites en entretien individuel et 53% de toutes les idées produites en *focus group* (cf. aussi tableau 8 ci-dessous). En revanche, aucune unité sémantique n'est du niveau de la dimension dans les entretiens individuels alors que c'est le cas de 6% d'unités sémantiques en *focus group*. Les *focus groups* jouissent donc d'une supériorité sur ce chapitre là.

De plus, le résultat concernant les idées nouvelles, qui sont aussi des idées détaillées, s'élève à 93% de toutes les idées produites en entretien individuel et à 64% de toutes les idées produites en *focus group* ($\chi^2= 17,3$ et $\rho<0.001$; cf. tableau 8 ci-dessous).

Le tableau 8 ci-après récapitule tous les résultats abordés jusque là ainsi que les quantités d'idées spécifiques et difficiles abordées dans les paragraphes suivants.

Tableau 8 : Comparaison de la nature des idées produites en entretien individuel versus en focus group

Type de technique	EI	EI (idées non ...)	FG	FG (idées non ...)
Idées nouvelles	52%	48%	37%	63%
Idées détaillées	79%	21%	53%	47%
Idées nouvelles détaillées ¹⁰¹	93%	7%	64%	36%
Idées spécifiques	55%	45%	13%	87%
Idées difficiles	10%	90%	34%	66%
Nombre total d'unités sémantiques	329		141	
Durées des entretiens	24h		5h	

4.2.3. Quantité d'idées spécifiques à chaque technique

Le pourcentage d'idées spécifiques à chaque technique a été calculé. Les explications apportées uniquement par les *focus groups* représentent 13% du total de leurs idées et les explications apportées uniquement par les entretiens individuels 55% du total de leurs idées ($\chi^2 = 68,75$, $p < 0,001$; cf. tableau 8).

De plus, une analyse supplémentaire met au jour que 84% des idées spécifiques aux *focus groups* sont des explications qui ont émergé au cours de discussions de confrontation où les participants ne partagent pas les mêmes points de vue (usages, motivations différentes, etc.) et se les explicitent. Les 16% restants correspondent à des explications qui ont été données par un participant dans une sorte de monologue s'adressant au groupe ou à l'animateur.

Concernant ces aspects divergents, certains participants désirent, par exemple, la possibilité d'une recherche par orthographe approchée de façon à trouver des CDs dont ils ne savent plus correctement orthographier le nom de l'artiste, tandis que d'autres redoutent cette solution, de crainte d'obtenir une liste trop exhaustive de résultats. Malgré le nombre élevé d'explications n'ayant été formulées que lors des entretiens individuels, cet aspect de la recherche, présentant un caractère important, a été évoqué uniquement dans les *focus groups*.

4.2.4. Méga-dimensions ou points essentiels générés par chaque technique

Dans ce paragraphe, nous poursuivons la comparaison des deux techniques concernant la nature du contenu verbal produit mais à un niveau « macro », ce qui permet de creuser les avantages que peuvent présenter les *focus groups*. Nous comparons les points essentiels dégagés dans chaque technique, c'est-à-dire les dimensions d'emblée collectives en *focus*

group et les dimensions individuelles, partagées par les interviewés en face à face. Elles sont présentées dans le tableau 9 ci-dessous. Les dimensions individuelles partagées ont pu être hiérarchisées par ordre d'importance (les plus importantes sont celles partagées par huit sujets et les moins importantes sont celles partagées par quatre sujets seulement).

Tableau 9 : Points essentiels dégagés dans les focus groups versus dans les entretiens individuels (équivalant aux dimensions collectives versus les dimensions individuelles partagées)

Dimensions collectives des <i>focus groups</i> (synthèse des titres des regroupements)	Dimension les plus partagées par les interviewés en face à face
<ul style="list-style-type: none"> • Notoriété, • Sécurité de paiement, • Confidentialité, sécurité de l'identité, • Page d'accueil, crédibilité du site et lisibilité du vendeur (carte de visite, mode d'emploi de lecture), • Esthétique, présentation, attractivité (mise en page, lisibilité), • Rapidité de chargement, accès technique (ADSL, ...), • Navigation, • Conseils d'achat et détails sur les produits, • Offre (diversité des marques, gammes de produits), • Services après-vente (dont la livraison). 	<ul style="list-style-type: none"> • Facilité du processus de recherche, • Description des produits, • Rapidité et facilité de navigation, • Pertinence et présentation des informations (en général), • Existence de différentes fonctions de recherche, • Disponibilité-livraison des produits, • Facilité du processus de commande, • Diversité et qualité des produits.

Les deux *focus groups* ont permis d'obtenir rapidement une synthèse de dix dimensions collectives, tandis que les huit dimensions individuelles partagées ont été obtenues après un considérable travail d'analyse et de codage des données. Les *focus groups* présentent, en ce sens, un avantage en termes de concision.

Les dimensions individuelles partagées portent sur les aspects d'utilisabilité du site Internet (excepté les dimensions en sixième et huitième positions) tandis que les dimensions collectives portent sur l'ensemble des aspects relatifs à un achat par Internet et davantage sur les aspects commerciaux (la notoriété du site, les aspects sécurité et confidentialité, l'intérêt de l'offre de produits, etc.).

Nous rappelons que la phase pour obtenir les dimensions les plus importantes n'a pas été réalisée (c'est-à-dire l'exercice de hiérarchisation des dimensions selon leur degré d'importance par un large échantillon du seul groupe d'utilisateurs à étudier lors de l'évaluation finale, cf. étape 3 du processus d'évaluation). Ces deux listes de méga-dimensions ne représentent ni l'une ni l'autre les dimensions les plus importantes, mais une synthèse des dimensions collectives/partagées de chaque technique.

4.2.5. Quantité d'idées difficiles à interpréter

Pour comparer la facilité d'exploitation de chaque technique, le pourcentage d'unités sémantiques difficiles à interpréter a été calculé pour chacune. Les *focus groups* comptent 34% du total de leurs idées qui sont des unités sémantiques un peu vagues et qui n'ont pu être comprises que grâce à une solide connaissance du domaine ou grâce aux entretiens individuels. Ce score est significativement plus élevé que celui des entretiens individuels qui n'en comptent que 10% ($\chi^2= 35.88$, $p<0,001$; cf. tableau 8 ci-dessus). La production des entretiens individuels est donc plus facile à exploiter que celle du focus group.

4.2.6. Les performances de la situation de verbalisation sans versus avec utilisation de sites marchands

Enfin, nous présentons les performances de la situation de verbalisation sans versus avec utilisation de sites marchands, dans chaque technique. Les pourcentages d'unités sémantiques nouvelles, détaillées, spécifiques et difficiles ont été calculés (par rapport au volume total d'unités de chaque situation de chaque technique). En réalité, seules les différences significatives sont considérées et résumées dans le tableau 10 ci-dessous.

Tableau 10 : Comparaison de la nature des idées produites dans la situation de verbalisation sans versus avec utilisation de sites marchands

Situation de verbalisation sans ou avec utilisation	Sans	Sans (idées non ...)	Avec	Avec (idées non ...)	Test de Chi-Carré
Idées nouvelles en FG	29%	71%	50%	50%	$\chi^2=6,47$; $p=0,0110$
Idées détaillées en EI	75%	25%	86%	14%	$\chi^2=6,33$; $p=0,0119$
Idées difficiles en EI	14%	86%	6%	94%	$\chi^2=5,26$; $p=0,0218$
Idées difficiles en FG	43%	54%	20%	80%	$\chi^2=7,29$; $p=0,007$

Légende : seules les différences significatives pour chaque technique (entretien individuel, focus group) sont présentées.

Les résultats révèlent que 50% des idées produites en situation de verbalisation avec utilisation lors des *focus groups* sont nouvelles. En revanche, seulement 29% des idées produites en situation sans utilisation en *focus group* sont nouvelles. La séquence d'utilisation ou de démonstration des sites marchands en *focus group* semble donc très propice à l'émergence de données nouvelles par rapport à la liste de dimensions experts.

Les verbalisations simultanées en entretien individuel favorisent légèrement, non pas spécifiquement l'émergence d'idées nouvelles (résultat positif mais non significatif sur ce

point), mais l'émergence d'idées détaillées et d'idées plus faciles à interpréter. En effet, la situation de verbalisation avec utilisation des entretiens individuels génère en proportion légèrement plus d'idées détaillées que leur situation sans utilisation (86% contre 75%) et légèrement moins d'idées difficiles à interpréter (6% contre 14%).

De même, la séquence de démonstration des *focus groups* contribue à produire (en proportion) moins d'unités sémantiques difficiles à interpréter par rapport à la situation sans utilisation (20% contre 43%).

Enfin, les différences entre l'une et l'autre situation en termes de quantité d'idées spécifiques ne sont pas significatives, et cela pour les deux techniques.

Après avoir caractérisé la nature et l'exploitabilité de la production verbale de l'entretien individuel et du *focus group*, puis des situations de verbalisation sans versus avec utilisation, nous en étudions la productivité. Nous réalisons cela uniquement pour chaque technique (les situations sans versus avec utilisation ne sont plus considérées).

5. DONNEES ET RESULTATS SUR LA PRODUCTIVITE DES TECHNIQUES

Dans ce paragraphe, la productivité en termes de ressources humaines et de coûts financiers est d'abord traitée, puis l'effet de saturation de cette productivité d'un entretien à un autre. Pour une meilleure compréhension, nous présentons, ensemble, l'analyse des données et les résultats pour chacun de ces deux points.

5.1. LA PRODUCTIVITE DE CHAQUE TECHNIQUE

La productivité en termes de coûts correspond à la quantité de données produites, divisée par le coût financier de réalisation d'une technique (préparation, recrutement, passation, etc.). La productivité en termes de ressources humaines correspond à la quantité de données produites, divisée par l'effort humain mobilisé (nombre de professionnels ou de participants, multiplié par leur(s) durée(s) d'intervention).

Pour se placer dans une situation réaliste, nous avons mesuré les ressources humaines, les différentes durées nécessaires et tous les frais qu'occasionnent la réalisation de ces entretiens, en nous basant sur des données recueillies auprès des professionnels de l'évaluation de France Télécom, mais aussi auprès de professionnels de cabinets d'évaluation indépendants¹⁰².

Nous examinons d'abord la productivité du nombre total d'idées produites par chaque technique (donc l'ensemble des idées quel que soit le niveau de détail auquel elles appartiennent). Dans un second temps, nous approfondissons en considérant la productivité du nombre total de dimensions, de sous-dimensions et de détails révélés dans chaque technique.

5.1.1. La productivité du nombre total d'idées

Les résultats concernant la productivité en termes de ressources humaines sont d'abord explicités, puis ceux concernant la productivité en termes de coûts financiers.

5.1.1.1. La productivité du nombre total d'idées en termes de ressources humaines

Dans les graphiques suivants (cf. figures 4 et 5), nous avons calculé le rapport de la productivité moyenne d'un *focus group* (avec huit participants) sur celle de huit entretiens individuels, concernant les ressources humaines mobilisées. La productivité est ici le nombre total d'idées divisé par la quantité de ressources humaines. Le résultat en ordonnée, dans les deux graphiques suivants, peut aussi s'exprimer de la façon suivante :

$$Y = \frac{\frac{\text{nombre total d'idées produites dans 1 FG}}{\text{nombre d'hommes/heure nécessaire à 1FG}}}{\frac{\text{nombre total d'idées produites dans 8EI}}{\text{nombre d'hommes/heure nécessaire à 8EI}}}$$

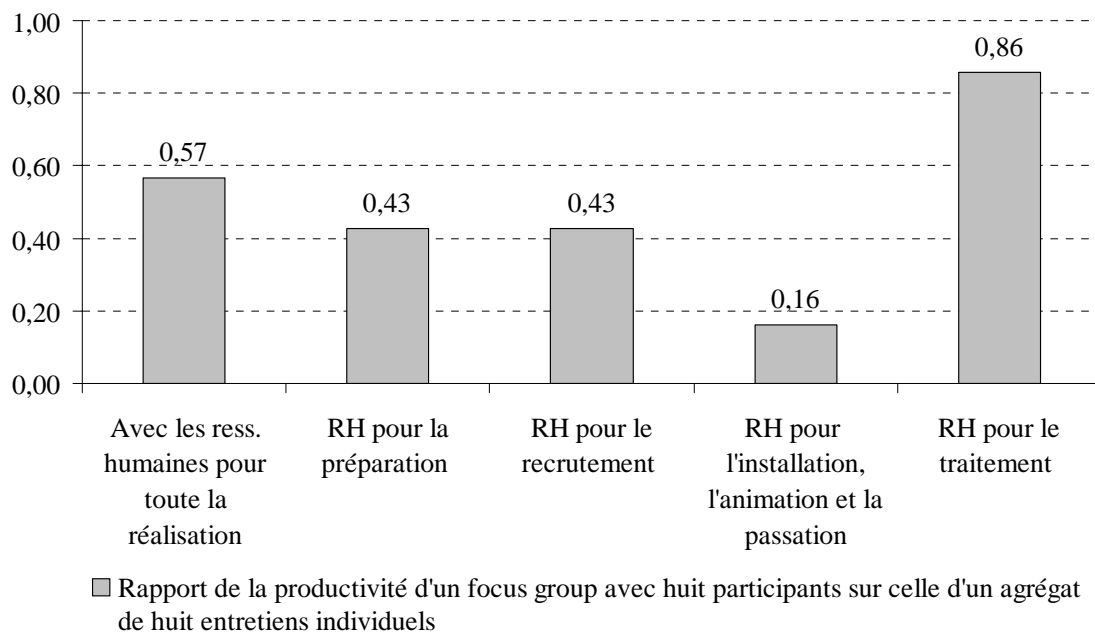
Les ressources humaines, en hommes/heure¹⁰³, sont le nombre de professionnels ou de participants recrutés multipliés par les différentes durées de réalisation (préparation, recrutement, passation, etc.). Le décompte de l'effort humain mobilisé ne tient pas compte de la rémunération des professionnels ou participants.

Dans le graphique ci-dessous (figure 4), nous avons considéré les rapports de productivités en termes de ressources humaines pour ¹⁰⁴ :

- toute la réalisation d'un *focus group* et de huit entretiens individuels,
- la préparation uniquement (un évaluateur),
- le recrutement uniquement (un « recruteur »),

- l'installation (un technicien), l'animation (un ou deux évaluateurs et le technicien) et la passation (un participant pour chaque entretien individuel, huit participants pour chaque *focus group*),
- et le traitement des données (un évaluateur).

Figure 4 : Comparaison du focus group et de l'entretien individuel concernant la productivité des idées en termes de ressources humaines, pour différentes phases de leur réalisation



Légende : il s'agit d'un rapport de productivités : si le résultat en ordonnée est supérieur à 1, le focus group est plus performant que l'agrégat d'entretiens individuels, tandis que s'il est inférieur à 1, c'est l'inverse. Par ailleurs, il n'y a pas d'unité précisée sur l'échelle des ordonnées car il s'agit du rapport de deux productivités.

Concernant la lecture du graphique ci-dessus, si le résultat en ordonnée est supérieur à 1, le *focus group* avec huit participants est plus performant que l'agrégat de huit entretiens individuels, tandis que s'il est inférieur à 1, c'est l'inverse. Ainsi, les résultats démontrent que la productivité de l'agrégat d'entretiens individuels est meilleure que celle du *focus group* concernant l'effort humain mobilisé pour la réalisation dans son ensemble comme pour les différentes phases de réalisation.

En comparaison à sa productivité globale, la productivité du *focus group* chute légèrement concernant la préparation et concernant le recrutement. Néanmoins, la productivité relative à la préparation est équivalente à celle relative au recrutement. En effet, dans les deux cas, les mêmes ressources humaines sont nécessaires que l'évaluation se réalise avec un *focus group* ou des entretiens individuels. Il faut le même effort pour documenter l'étude et préparer le guide d'entretien ou pour recruter le même nombre de participants¹⁰⁵.

Ainsi pour la productivité relative à la préparation comme pour celle relative au recrutement, on trouve une relation d'équivalence à : $1FG \approx 3 \frac{1}{2} EI$ ¹⁰⁶.

La productivité du *focus group* chute bien davantage concernant l'effort mobilisé pour l'installation/animation/passation ($1FG \approx 1EI$), essentiellement en raison des huit participants mobilisés simultanément (huit hommes/heure) au lieu d'un participant pour chaque entretien individuel (un homme/heure)¹⁰⁷.

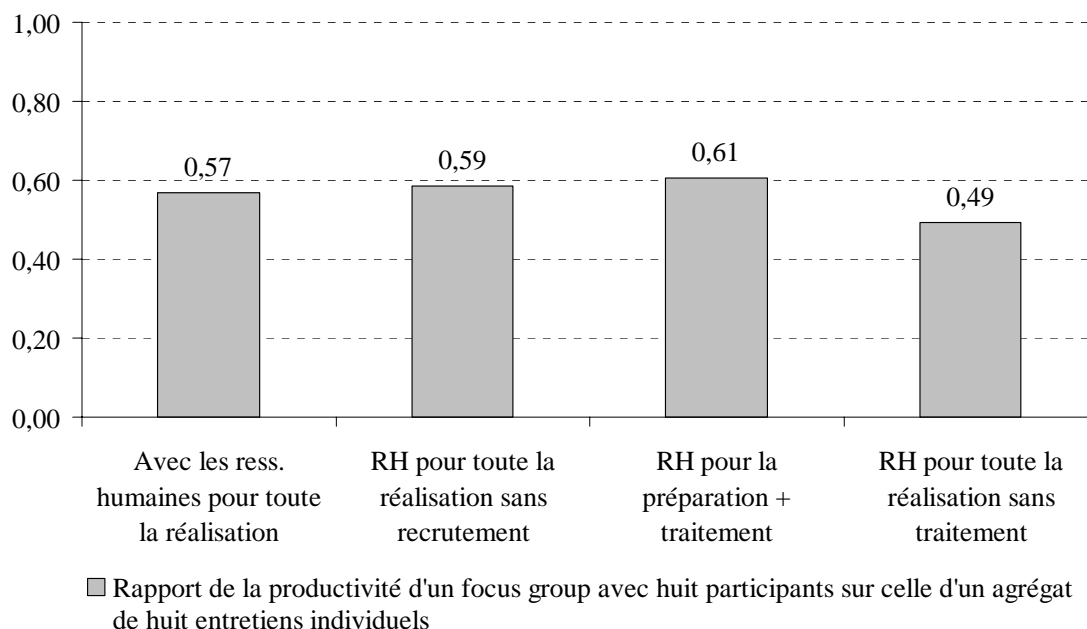
En revanche, la productivité du *focus group* s'améliore nettement concernant l'effort mobilisé pour le traitement des données. Le *focus group* est presque aussi performant que l'agrégat d'entretiens individuels concernant ce point ($1FG \approx 7EI$). D'après les professionnels de France Télécom, comme d'après ceux des cabinets d'évaluation, le temps de traitement des données est équivalent à deux fois le temps de passation pour les entretiens individuels et à trois fois le temps de passation pour le *focus group* (sans la rédaction du rapport¹⁰⁸). Si nous comptons seulement deux fois le temps de passation pour le traitement du *focus group* (ce qui est très téméraire selon les professionnels), la productivité du *focus group* est alors équivalente à celle de l'agrégat d'entretiens individuels ($1FG = 8EI$).

Les précédents résultats donnent des indications pour chaque étape considérée individuellement. Le graphique ci-après (cf. figure 5) expose des résultats plus opérationnels. En effet, différents scénarios de réalisation d'une évaluation sont envisagés, correspondant à des indications données par les professionnels de l'évaluation. Les quatre résultats du graphique correspondent à :

- toute la réalisation d'un *focus group* et de huit entretiens individuels,
- puis toute la réalisation sans le recrutement des participants, car le service d'évaluation (et même certains cabinets d'évaluation) a ses propres ressources de recrutement. Ce résultat permet donc d'observer la productivité en écartant cette ressource,
- puis la réalisation sans le recrutement et sans la passation/animation ce qui équivaut aux seules ressources de préparation de l'évaluation et de traitement des données. Ce résultat est un indicateur dans le cas où le service d'évaluation sous-traite l'évaluation, et le responsable de l'évaluation ne fait alors que la préparer et en traiter les données,
- et enfin, toute la réalisation sans le traitement des données : il s'agit d'un résultat demandé par le service d'évaluation car il permet d'isoler la part relative que

représentent les ressources de traitement des données, même s'il est plus rare que l'analyse des données soit sous-traitée.

Figure 5 : Comparaison du focus group et de l'entretien individuel concernant la productivité des idées en termes de ressources humaines, pour différents scénarios de réalisation



Légende : il s'agit d'un rapport de productivités : si le résultat en ordonnée est supérieur à 1, le focus group est plus performant que l'agrégat d'entretiens individuels, tandis que s'il est inférieur à 1, c'est l'inverse. Par ailleurs, il n'y a pas d'unité précisée sur l'échelle des ordonnées car il s'agit du rapport de deux productivités

Premièrement, le graphique montre que les rapports de productivités en termes de ressources humaines sont très similaires, concernant la réalisation complète et la réalisation sans le recrutement. Le recrutement représente donc un poids relatif négligeable par rapport aux autres ressources humaines engagées. Cela confirme les précédents résultats (seulement une baisse légère de la productivité avec le seul recrutement considéré par rapport à la productivité globale).

Deuxièmement, un rapport de productivités similaire est aussi obtenu si le service d'évaluation sous-traite l'évaluation (cf. résultat « préparation + traitement »). Seul l'effort humain fourni en interne concernant la préparation et le traitement est pris en compte dans ce calcul (l'effort humain mobilisé pour l'évaluation sous-traitée ne l'est pas). Concernant cet effort humain fourni en interne, le *focus group* présente la même faible productivité si le service sous-traite l'évaluation ou la réalise en entier lui-même (ou ne sous-traite que le recrutement). La productivité d'un agrégat d'entretiens individuels est alors presque deux fois supérieure à celle du *focus group* ($8EI \approx 1\frac{1}{2}FG$ ou $1FG \approx 5EI$).

Le dernier résultat confirme que le traitement des données constitue le point fort du *focus group* par rapport aux autres phases de réalisation (sans pour autant devenir plus performant que l'entretien individuel). Sans la prise en compte de la phase de traitement (où de plus petites ressources sont nécessaires au *focus group*¹⁰⁹), la productivité du *focus group* est légèrement moins bonne par rapport aux trois autres résultats. Néanmoins, dans le rare cas où le service sous-traite l'analyse des données d'une évaluation, la productivité du *focus group* est seulement légèrement diminuée (1FG≈4EI).

5.1.1.2. La productivité du nombre total d'idées en termes de coûts financiers

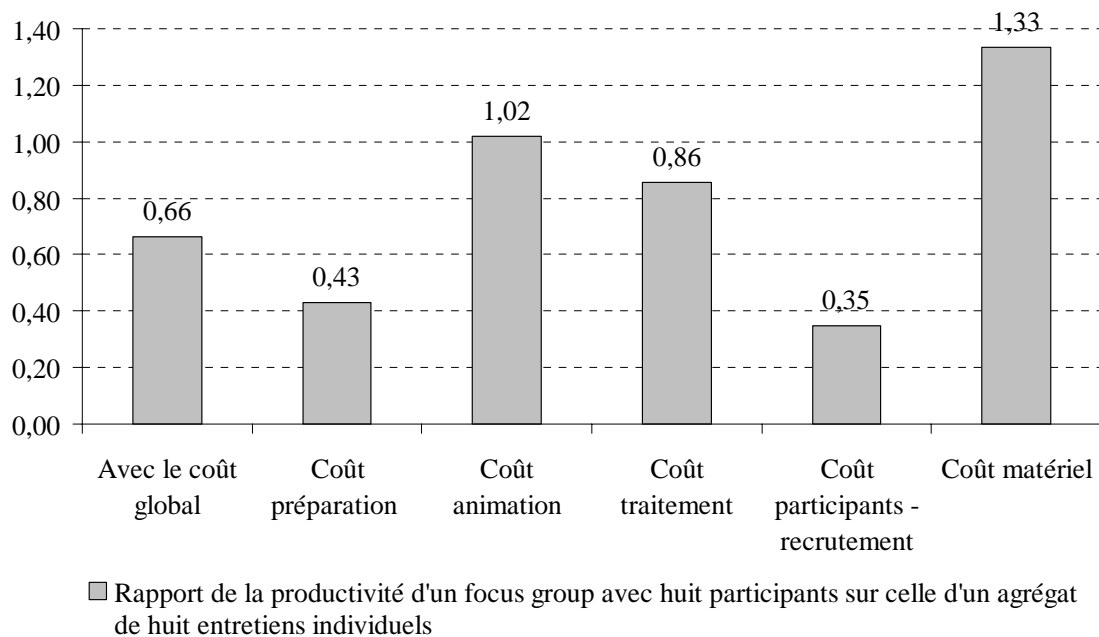
Dans les graphiques suivants (cf. figures 6 et 7), le rapport de la productivité moyenne d'un *focus group* (avec huit participants) sur celle de huit entretiens individuels a été calculé, concernant les coûts financiers engagés. La productivité, dans ce cas, est le nombre total d'idées divisé par les dépenses en Euros. Le résultat en ordonnée, dans les deux graphiques suivants, peut s'exprimer de la façon suivante :

$$Y = \frac{\frac{\text{nombre total d'idées produites dans 1 FG}}{\text{nombre d'Euros dépensés pour 1FG}}}{\frac{\text{nombre total d'idées produites dans 8EI}}{\text{nombre d'Euros dépensés pour 8EI}}}$$

Dans le graphique ci-dessous (figure 6), nous avons considéré les rapports de productivités en termes de dépenses¹¹⁰ pour :

- toute la réalisation d'un *focus group* et de huit entretiens individuels,
- la préparation (rémunération d'un évaluateur),
- l'animation (rémunération d'un ou deux évaluateurs et leur « snack » lors des *focus groups*),
- le traitement des données (rémunération d'un évaluateur),
- le coût des participants (rémunération d'un « recruteur » pour leur recrutement, indemnités des participants et « snacks » lors des *focus groups*¹¹¹),
- le coût matériel (location de la salle équipée, comprenant les services du technicien).

Figure6 : Comparaison du focus group et de l'entretien individuel concernant la productivité des idées en termes de coûts financiers, pour différentes phases de leur réalisation



Légende : il s'agit d'un rapport de productivités : si le résultat en ordonnée est supérieur à 1, le focus group est plus performant que l'agrégat d'entretiens individuels, tandis que s'il est inférieur à 1, c'est l'inverse. Par ailleurs, il n'y a pas d'unité précisée sur l'échelle des ordonnées car il s'agit du rapport de deux productivités

Ces résultats montrent que la productivité de l'entretien individuel est meilleure que celle du *focus group* concernant les coûts engagés pour toute la réalisation comme pour plusieurs phases ou coûts de réalisation, excepté pour les frais d'animation des entretiens et le coût matériel.

En effet, concernant l'animation, seuls deux évaluateurs doivent être rémunérés pendant la soirée du *focus group* (ainsi que leur « snack ») contre un évaluateur pendant deux jours et demi pour les huit entretiens individuels. Néanmoins, en raison de la moindre production verbale du *focus group* par rapport à huit entretiens individuels, la productivité du *focus group* est seulement équivalente (et non supérieure) à celle de l'agrégat d'entretiens individuels (1FG=8EI).

Concernant les aspects matériels, le coût du *focus group* est très nettement moindre par rapport à celui des entretiens individuels (location d'une salle pendant une journée pour la soirée du *focus group* contre trois jours et demi pour les entretiens individuels avec leur installation). Ainsi, la productivité en termes de coûts du *focus group* devient supérieure à celle des entretiens individuels concernant ce point (1FG≈11EI).

Par ailleurs, le rapport de productivités en termes de coûts est le même que celui en termes de ressources humaines (cf. figure 4, pages précédentes) concernant la préparation de l'évaluation. De même que précédemment, les coûts financiers pour préparer une évaluation

sont les mêmes, qu'elle se réalise avec huit entretiens individuels ou avec un *focus group* de huit participants. La productivité du *focus group* est ainsi toujours relativement faible concernant la préparation (FG \approx à plus de 3 EI).

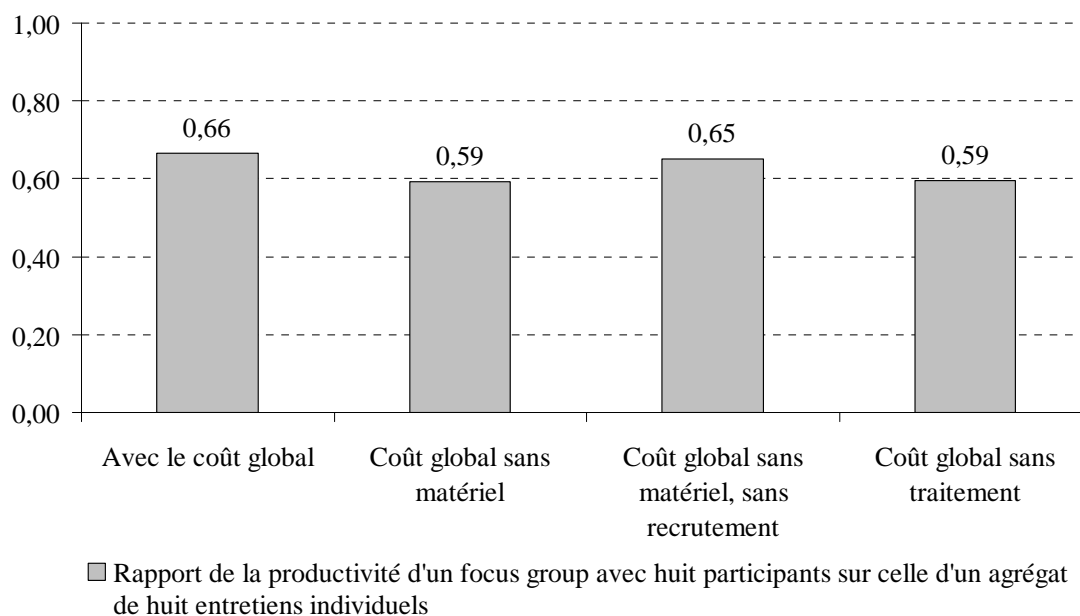
Le rapport de productivités en termes de coûts est aussi le même que celui en termes de ressources humaines concernant le traitement des données (cf. figure 4). Cependant, dans ce cas, le coût de traitement du *focus group* est moindre par rapport à celui des entretiens individuels. Ainsi, la productivité en termes de coûts, relative au traitement du *focus group*, est presque équivalente à celle des entretiens individuels (1FG \approx 7EI) et le serait tout à fait si un traitement plus rapide pouvait être mis en œuvre (deux au lieu de trois fois le temps de passation).

Enfin, les frais relatifs aux participants et à leur recrutement pénalisent davantage la performance du *focus group* que celle des entretiens individuels en raison des indemnités plus importantes pour deux heures et demie de *focus group* que pour une heure et demie d'entretien individuel et des huit « snacks » à offrir. La rémunération du « recruteur », frais prépondérant par rapport aux précédents, sont les mêmes dans les deux cas.

De même que précédemment, les résultats ci-dessus donnent des indications pour chaque étape considérée individuellement. Le graphique ci-dessous (cf. figure 7) expose des résultats plus opérationnels correspondant aux différents scénarios de réalisation suivants :

- toute la réalisation d'un *focus group* et de huit entretiens individuels,
- toute la réalisation sans la location du matériel, puisque le service d'évaluation a son propre laboratoire équipé et ne paie donc pas les frais de location (ce qui n'est pas le cas de tous les cabinets d'évaluation),
- toute la réalisation sans la location du matériel et sans le recrutement des participants : il s'agit là d'un résultat plus réaliste car ces coûts sont transparents pour le service d'évaluation qui a aussi ses propres ressources de recrutement. Il ne se facture pas en interne ces coûts mais les évalue pour les facturer au client de l'évaluation (de même, certains cabinets disposent de ressources matérielles et de recrutement, et minimisent alors ces frais pour proposer un coût global d'étude. Cela n'est pas le cas si seul le recrutement des participants leur est commandité),
- et enfin, toute la réalisation sans le traitement des données : de même que précédemment, il s'agit d'un résultat davantage à titre indicatif.

Figure 7 : Comparaison du focus group et de l'entretien individuel concernant la productivité des idées en termes de coûts financiers, pour différents scénarios de réalisation



Légende : il s'agit d'un rapport de productivités : si le résultat en ordonnée est supérieur à 1, le focus group est plus performant que l'agrégat d'entretiens individuels, tandis que s'il est inférieur à 1, c'est l'inverse. Par ailleurs, il n'y a pas d'unité précisée sur l'échelle des ordonnées car il s'agit du rapport de deux productivités

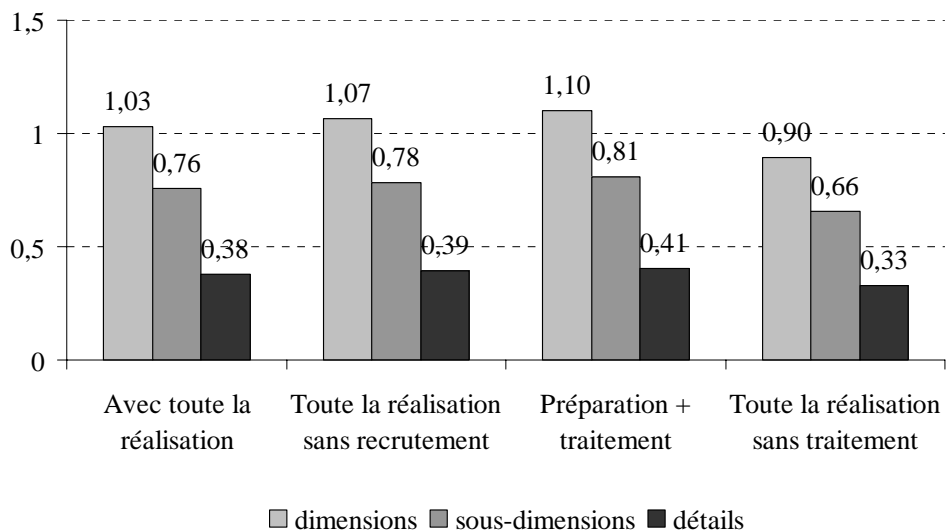
Selon le graphique ci-dessus, les rapports de productivités en termes de coûts concernant les quatre scénarios détaillés ci-dessus sont très similaires. Quel que soit le cas de figure ou la distinction frais internes/externes envisagés par le service d'évaluation, la productivité d'un agrégat d'entretiens individuels est presque toujours deux fois supérieure à celle du *focus group* ($8EI \approx 1\frac{1}{2}FG$ ou $1FG \approx 5EI$). Dans ce qui précède, les différentes productivités en termes de coûts considérés individuellement sont donc à relativiser. Par exemple, la plus grande productivité du *focus group* en termes de coût matériel est certes réelle, mais négligeable par rapport à la productivité globale ou tout du moins insuffisante à la rehausser.

5.1.2. La productivité des dimensions utilisateurs, des sous-dimensions et des détails

Afin d'approfondir les précédents résultats, nous étudions la productivité du nombre total de dimensions utilisateurs, de sous-dimensions et de détails révélés dans chaque technique. Selon les résultats du paragraphe 4.2.2. (tableau 7), les *focus groups* semblent plus performants à révéler des dimensions (ou données d'un niveau de détail très général) que les entretiens individuels, et inversement concernant la production de données du niveau du détail. Nous cherchons à estimer ce qu'il en est précisément, en examinant la productivité de ces données.

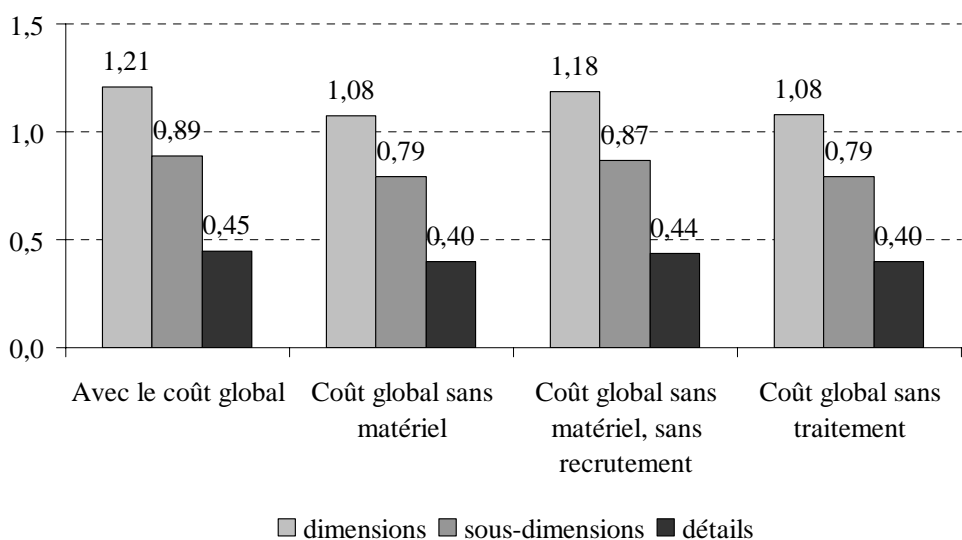
Les graphiques ci-après exposent d'abord le résultat concernant la productivité en termes de ressources humaines (cf. figure 8), puis celui concernant la productivité en termes de coûts financiers (cf. figure 9), et cela pour les mêmes scénarios de réalisation que précédemment.

Figure 8 : Comparaison du focus group et de l'entretien individuel concernant la productivité de données verbales de différents niveaux de détail en termes de ressources humaines (trois niveaux de détail : dimensions, sous-dimensions, détails)



Remarque : il s'agit d'un rapport de productivités : si le résultat en ordonnée est supérieur à 1, le focus group est plus performant que l'agrégat d'entretiens individuels, tandis que s'il est inférieur à 1, c'est l'inverse

Figure 9 : Comparaison du focus group et de l'entretien individuel concernant la productivité de données verbales de différents niveaux de détail en termes de coûts financiers (trois niveaux de détail : dimensions, sous-dimensions, détails)



Remarque : il s'agit d'un rapport de productivités : si le résultat en ordonnée est supérieur à 1, le focus group est plus performant que l'agrégat d'entretiens individuels, tandis que s'il est inférieur à 1, c'est l'inverse.

Globalement, nous retrouvons les mêmes petites variations que précédemment d'un scénario de réalisation à un autre, pour les deux types de productivité. Par ailleurs, les conclusions de la meilleure performance du *focus group* pour produire des données d'un niveau général sont confirmées de même que celle de l'entretien individuel pour produire des données détaillées. La productivité des dimensions est quasiment toujours meilleure en *focus group* qu'en entretien individuel (pour tous les résultats, une relation d'équivalence peut être trouvée à environ : $1FG \approx 8EI$ ou $1FG \approx 9EI$). C'est l'inverse concernant la production des détails ($1FG \approx 3EI$).

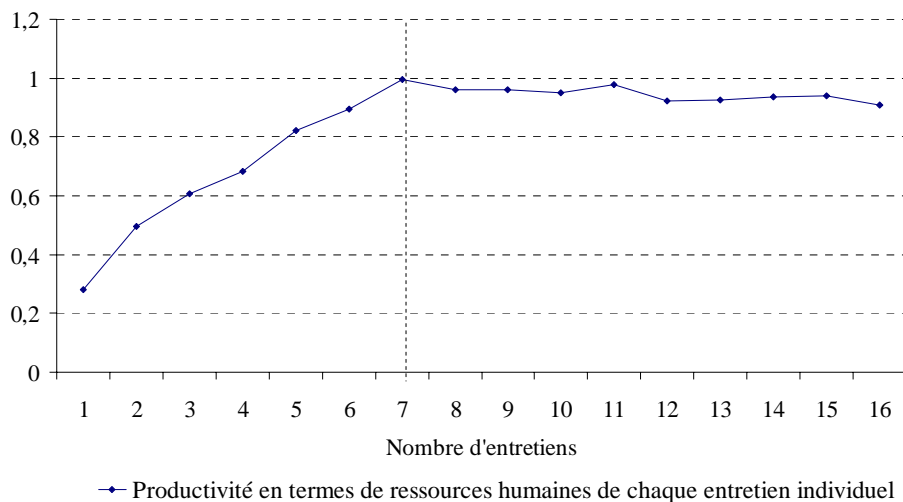
5.2. L'EFFET DE SATURATION DE LA PRODUCTIVITE DU NOMBRE TOTAL D'IDEES EN ENTRETIEN INDIVIDUEL

Pour renseigner l'intérêt de mener un entretien supplémentaire par rapport au(x) précédent(s) et donc le seuil à partir duquel la productivité stagne ou chute, il s'agit d'observer son évolution d'un entretien à un autre. En fait de huit entretiens individuels, 16 ont été menés avec chaque type de sujets (« plus experts » ou « plus novices »). Cela permet d'observer l'évolution de la productivité avec une quantité suffisamment intéressante. Nous rappelons que seules les idées nouvelles apportées à chaque entretien supplémentaire ont été conservées.

En revanche, n'ayant pu réaliser qu'un *focus group* avec chaque population de sujets, nous ne sommes pas en mesure de produire de tels résultats concernant la situation collective.

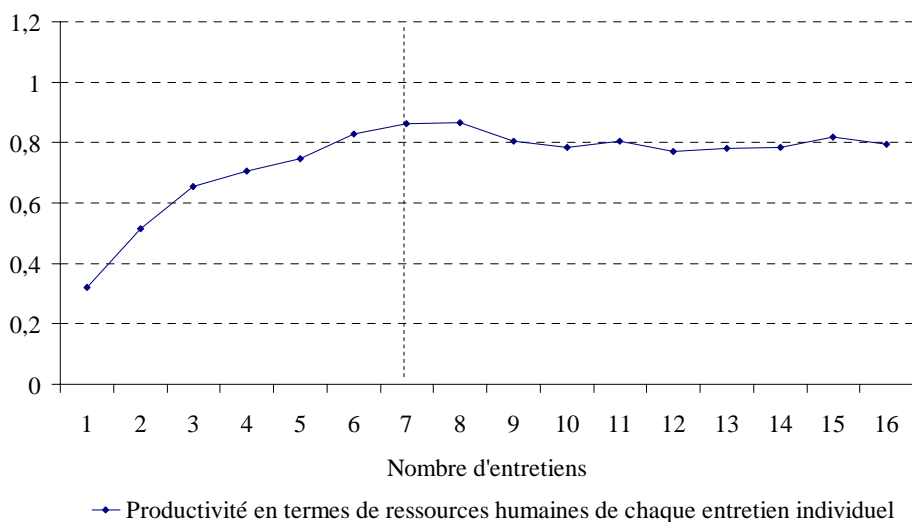
Les graphiques ci-après montrent l'évolution de la productivité du nombre total d'idées (donc l'ensemble des idées quel que soit le niveau de détail auquel elles appartiennent) d'un entretien individuel à un autre jusqu'au seizième. L'évolution de la productivité en termes de ressources humaines est d'abord présentée (cf. figures 10 et 11), puis celle en termes de coûts financiers (cf. figures 12 et 13), toutes deux, avec les 16 entretiens individuels avec les sujets « plus expérimentés », d'une part, et avec les 16 avec les sujets « plus novices », d'autre part.

Figure 10 : Evolution de la productivité des idées en termes de ressources humaines d'un entretien individuel à un autre, avec les sujets « plus expérimentés »



La productivité est exprimée en nombre d'idées par homme/heure.

Figure 11 : Evolution de la productivité des idées en termes de ressources humaines d'un entretien individuel à un autre, avec les sujets « plus novices »



La productivité est exprimée en nombre d'idées par homme/heure.

Figure 12 : Evolution de la productivité des idées en termes de coûts financiers d'un entretien individuel à un autre, avec les sujets « plus expérimentés »

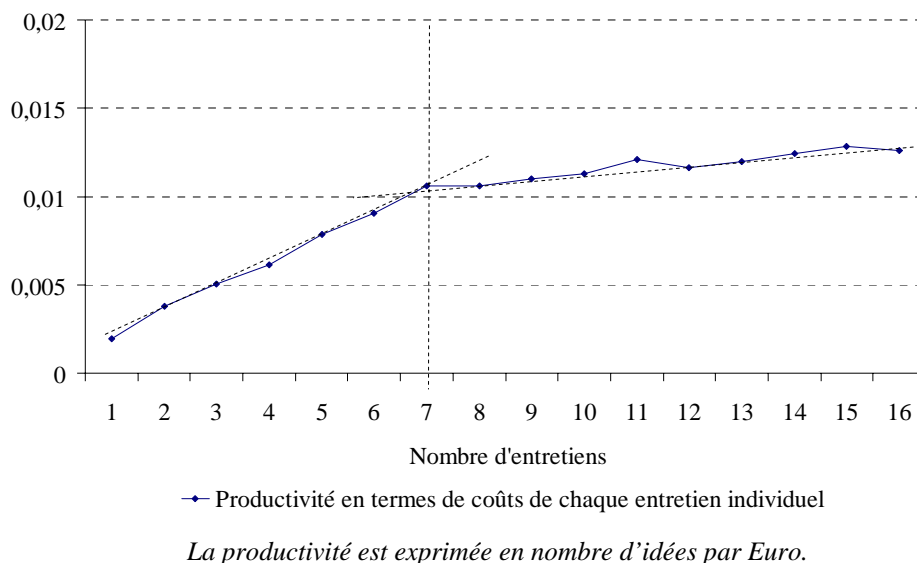
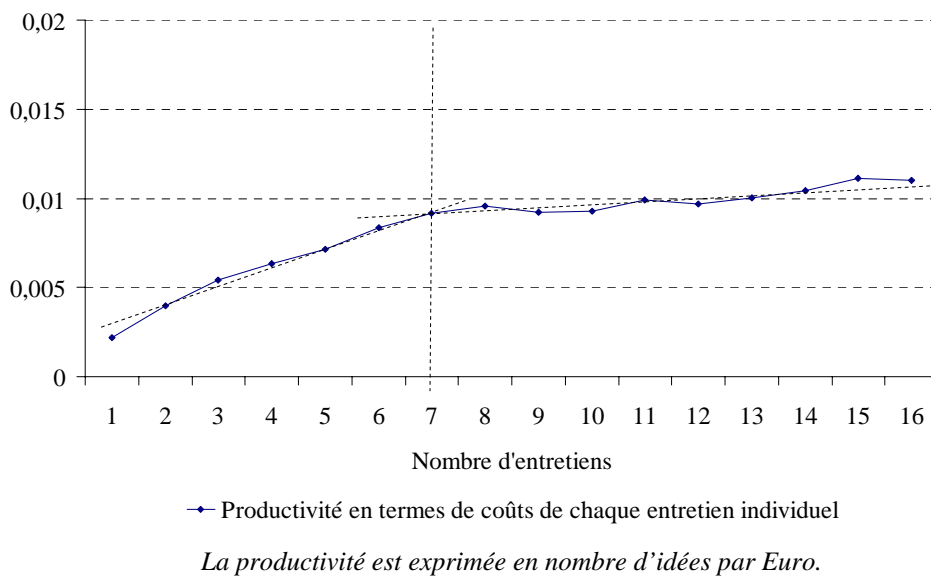


Figure 13 : Evolution de la productivité des idées en termes de coûts financiers d'un entretien individuel à un autre, avec les sujets « plus novices »



Ces résultats mettent au jour qu'après sept entretiens, la productivité en termes de ressources humaines augmente plus faiblement, voire diminue légèrement, et cela avec les sujets « plus expérimentés » comme avec les sujets « plus novices ». De même, après sept entretiens, la productivité en termes de coûts financiers augmente plus faiblement (pour les deux catégories de sujets). En d'autres termes, le coût humain ou financier de chaque idée

produite augmente sensiblement à partir du septième entretien. Il n'est plus aussi rentable de mener un entretien supplémentaire.

6. SYNTHÈSE DES RESULTATS

Le tableau ci-dessous synthétise les résultats des paragraphes 4 et 5.

Tableau 11 : Récapitulatif des résultats de comparaison de l'entretien individuel versus du focus group quant à la nature des données produites et à leur productivité

		EI	FG
Contenu	Apport et quantité d'idées nouvelles / données experts	OUI : +	OUI : -
	Quantité d'idées détaillées	+	-
	Quantité de dimensions (données d'un niveau de détail très général)	-	+
	Quantité d'idées spécifiques à chaque technique	+	-
	Nature des méga-dimensions	Utilisabilité du site Internet ; Obtenues après un codage systématique	Aspects relatifs à l'achat par Internet ; Obtenues très rapidement
Exploitabilité	Quantité d'idées faciles à interpréter	+	-
Verbalisations avec et sans utilisation	Avec utilisation du produit/service	++ d'idées détaillées	+ d'idées nouvelles
	Sans utilisation du produit/ service	+ d'idées difficiles à interpréter	+ d'idées difficiles à interpréter
Productivité	Productivité du nombre total d'idées / ressources humaines : • pour la préparation, • pour le recrutement, • pour l'animation/passation, • pour le traitement.	+ + ++ presque ≈	- - -- presque ≈
	Productivité globale du nombre total d'idées / ressources humaines :	La productivité globale de 8EI est presque deux fois supérieure à celle du FG, que le service d'évaluation réalise l'évaluation en entier, sous-traite sa réalisation ou son analyse.	
	Productivité du nombre total d'idées / coûts financiers : • pour la préparation, • pour l'animation, • pour le traitement, • concernant le coût relatif aux participants + recrutement, • concernant le coût matériel.	+ équivalent presque ≈ + -	- équivalent presque ≈ - +
	Productivité globale du nombre total d'idées / coûts financiers :	8EI ≈ 1½FG ou 1FG ≈ 5EI, que le service d'évaluation considère ou non les frais de recrutement/matériel, voire sous-traite l'analyse des données.	
	Productivité des dimensions / ressources humaines ou coûts	- 8EI ≤ 1FG	+ 1FG ≈ 8 à 9EI
	Productivité des détails / ressources humaines ou coûts	+ 8EI ≈ 2 ½ FG	- 1FG ≈ 3EI
	Effet de saturation d'un EI à un autre	Evolutions de la productivité du nombre total d'idées / ressources humaines	Après 7EI, la productivité augmente plus faiblement, voire diminue.
Evolution de la productivité du nombre total d'idées / coûts		Après 7EI, la productivité augmente plus faiblement.	

7. PERFORMANCES DES DEUX TECHNIQUES UTILISEES EN VUE DE LA CONSTRUCTION D'UN QUESTIONNAIRE

Avant de dégager les conclusions, une remarque préliminaire s'impose : le traitement et codage des données que nous avons réalisés ne sont pas ceux qu'auraient effectués les professionnels de l'évaluation. Ces derniers, pour des raisons de temps, se basent généralement sur leurs notes, sur les éléments écrits par les sujets et, parfois, sur la visualisation des vidéos. Ils font très rarement retranscrire les entretiens. En ce sens, notre codage, beaucoup plus fin, n'est pas tout à fait représentatif, mais les professionnels de l'évaluation poursuivent le même objectif : mettre en évidence les dimensions toutes différentes, relatives à la perception des utilisateurs.

Nous soulevons là la question de la validité. En se plaçant dans des conditions optimales de comparaison de l'entretien individuel et du *focus group*, quels sont les aspects du « monde réel » qui sont négligés ? Plus que de validité d'une étude ou de généralisation en tant que vérité universelle, Lincoln et Guba (1985) parlent de « transférabilité ». Cela réfère à la possibilité de transférer ce qui est appris à partir d'un contexte d'étude à une application ayant un autre contexte mais présentant un certain degré de congruence avec le contexte de départ. De même, du fait des contextes proches (même tâche de génération d'idées et finalité commune de production de dimensions), nous pensons que nos conclusions sont au moins partiellement transposables. Les professionnels de l'évaluation peuvent s'appuyer dessus en les considérant comme des orientations qu'ils pourront encore affiner au fur et à mesure de leurs études appliquées.

Nous présentons d'abord les conclusions qui concernent l'entretien individuel et le *focus group*, utilisés comme techniques d'élicitation d'idées pour l'élaboration de dimensions utilisateurs. Nous abordons ensuite les limites que présente cette étude et enfin élargissons nos réflexions à l'utilisation de ces techniques, surtout le *focus group*, pour la conception de produits/services.

7.1. INTERET DES DONNEES UTILISATEURS PAR RAPPORT AUX DONNEES EXPERTS

Les entretiens individuels comme les *focus groups* apportent une part non négligeable d'éléments nouveaux par rapport à la liste de dimensions élaborées par les experts, même si

une majorité d'idées nouvelles sont des éléments de détail (donc en quelque sorte de petites nouveautés).

Nous en déduisons que les données utilisateurs sont plus riches que les données experts, de même que les tests utilisateurs sont plus performants pour détecter un maximum de problèmes d'utilisabilité par rapport aux méthodes expertes d'inspection (Desurvire, 1994; Karat, 1994). En outre, les entretiens individuels le sont davantage que les *focus groups*. Ils présentent donc un intérêt certain qui est développé dans ce qui suit.

7.2. AVANTAGES DE L'ENTRETIEN INDIVIDUEL POUR L'ELABORATION DES DIMENSIONS UTILISATEURS

Nous développons les conclusions relatives aux résultats concernant la nature, l'exploitabilité et la productivité des données en entretien individuel, puis nous proposons une recommandation sur le nombre d'entretiens à mener. Nous évoquons aussi brièvement les performances des verbalisations simultanées versus « ultra-consécutives »¹¹².

7.2.1. Avantages concernant la nature et l'exploitabilité des idées produites

Cette étude révèle que les entretiens individuels produisent proportionnellement plus d'idées qui n'ont pas émergé en *focus group*, plus d'idées nouvelles, plus d'idées détaillées et enfin plus d'idées faciles à interpréter, dans le cadre d'une tâche de génération d'idées. Fern (1982) conclue en préconisant les entretiens individuels si l'objectif est de générer une longue liste d'idées et pensées concernant un concept relativement complexe (comme la perception). Nous déduisons, de même, que la situation individuelle est la plus favorable pour l'élaboration d'une liste exhaustive de dimensions de perception utilisateurs avec une grande précision. Les professionnels de l'évaluation étaient parvenus au même avis favorable pour les entretiens individuels dans une précédente étude sur les sites Internet portail (Pican, 2001).

Cette conclusion est d'autant plus réelle si les évaluateurs ont encore une connaissance pauvre du domaine. Une bonne connaissance du thème à étudier est indispensable à la réalisation d'un questionnaire (Ghiglione & Matalon, 1998; Javeau, 1982; Mucchielli, 1968). Or, dans les données détaillées, les attentes et critères de jugement des sujets sont bien décrits, et ces données sont nombreuses en entretien individuel.

Par ailleurs, les verbalisations libres simultanées favorisent la production de données détaillées et faciles à interpréter. Les verbalisations « ultra-consécutives », soit à partir de précédentes expériences vécues, semblent moins performantes. Il peut ainsi être envisagé d'accroître la séquence d'utilisation de produits/services en se basant sur la méthode des verbalisations libres.

7.2.2. Avantages concernant la productivité

En plus de leurs avantages en termes de nature et d'exploitabilité de leurs données, les entretiens individuels présentent également des avantages en termes de productivité. Cette étude montre que la productivité d'un agrégat de huit entretiens individuels est presque deux fois supérieure à celle du *focus group* pour différents scénarios de réalisation (évaluation réalisée en entier ; recrutement et passation sous-traités ; traitement des données sous-traité ; frais de recrutement/matériel comptabilisés ou non), et cela qu'il s'agisse de l'effort humain ou des dépenses engagés. Ce résultat porte sur la productivité du nombre total d'idées, il est encore plus convaincant concernant celle des idées détaillées. La productivité d'un agrégat de huit entretiens individuels est alors plus de deux fois supérieure à celle du *focus group*.

Il semble que le seul point fort relatif du *focus group* soit sa productivité en termes de coût matériel en raison du temps très court de location d'une salle. Son plus grand point faible semble être sa productivité en termes de ressources humaines mobilisées pour l'animation/passation. Tous ces résultats vont à l'encontre de ce qui est communément accepté puisque le *focus group* est plus souvent reconnu comme une méthode très efficace. Cette conclusion est généralement spéculée sans considérer toutes les ressources humaines mobilisées ou les coûts de façon précise (cf. chapitre 2). Le *focus group* peut alors apparaître performant quant à la passation (concentration de données avec une courte durée d'entretien). De même, nos *focus groups* sont plus rapides dans l'absolu, puisque cinq heures de *focus group* sont comparés à 24 heures d'entretiens individuels (et de nombreuses dimensions sont révélées à un niveau général), mais ils n'apparaissent pas plus efficaces.

Morgan et Krueger (1993) tiennent, de même, pour mythique cette idée selon laquelle le *focus group* est plus rapide et moins cher et affirment que pour préparer un bon guide de *focus group*, il faut du temps et qu'un *focus group* sous-traité coûte cher. Ce coût est estimé de 1000 à 4000\$ (Morgan & Krueger, 1993) et peut même atteindre 6800\$ (aux tarifs de New York ; Nielsen, 1998). De même, Griffin et Hauser (1993) ne font pas de distinction entre

l'entretien en face à face et le *focus group* dans leur estimation du coût minimum d'un entretien pour éliciter les besoins utilisateurs/clients (1000 à 2000\$).

7.2.3. Recommandation sur le nombre de sujets

Enfin, du point de vue de la productivité, il semble inutile de recruter plus de sept sujets. Au-delà de sept entretiens individuels, l'information devient plus coûteuse (sur le plan humain comme financier), même si elle n'est pas nécessairement inutile. Cela est vrai avec les sujets « plus experts » comme avec les sujets « plus novices ». De plus, ces résultats ne sont pas spécifiques à un seul site marchand, puisque huit sites marchands, en tout, ont été visités. Toutefois, les tâches expérimentales concernaient l'exploration de la page d'accueil et la recherche d'un produit jusqu'au remplissage des bons de commande et non un achat réel. Avec une tâche plus réaliste encore, peut-être les résultats seraient-ils différents ?

Par ailleurs, il n'est pas mis en évidence le pourcentage d'une quantité, définie comme référence, de critères de jugement révélés par ces sept entretiens¹¹³. Il n'est pas possible de faire l'analogie avec les recommandations selon lesquelles cinq tests utilisateurs suffisent ou que 20 tests utilisateurs détectent la quasi totalité des problèmes d'utilisabilité d'un produit/système ou 30 entretiens individuels la quasi totalité des besoins utilisateurs/client s¹¹⁴. D'autres recherches mettant en évidence le pourcentage de critères révélés et non la productivité des idées d'un entretien à un autre nous semblent nécessaires pour cela (en travaillant à partir d'une liste simple et arrêtée de critères ou dimensions utilisateurs - sans les trois niveaux de détail).

Après avoir examiné les avantages de l'entretien individuel, nous étudions ceux du *focus group*, moins évidents mais néanmoins existants.

7.3. AVANTAGES DU FOCUS GROUP POUR LA CONSTRUCTION D'UN QUESTIONNAIRE DE TYPE « GENERAL »

Nous nous attachons ici à apporter des éléments de connaissances très opérationnels, utiles directement aux professionnels du service d'évaluation, car en rapport avec le processus qu'ils ont élaboré. Nous cherchons à répondre à la question suivante : le *focus group* serait-il un moyen de réduire ce processus d'évaluation ? D'autre part, sur un plan plus général, nous examinons quel type de questionnaire pourrait être construit avec les méga-dimensions produites en *focus group* et avec celles produites en entretien individuel.

7.3.1. Le focus group, une solution pour raccourcir le processus d'évaluation ?

Le processus d'évaluation pourrait-il être raccourci grâce aux méga-dimensions obtenues rapidement, sur la base des constructions collectives des *focus groups* ? Nous rappelons que les professionnels de l'évaluation génèrent une liste exhaustive de dimensions à partir d'entretiens individuels avec plusieurs groupes d'utilisateurs pour la faire réduire ensuite grâce à un travail de hiérarchisation par un large échantillon d'utilisateurs spécifiques. Les dimensions les plus importantes ainsi obtenues sont celles à partir desquelles ils réalisent le questionnaire. Pour apporter des éléments de réponse à cette question, nous analysons plus en détail l'étude sur les sites Internet portail, déjà évoquée et similaire à celle-ci.

Lors de cette étude, douze entretiens individuels sont comparés à deux *focus groups*. Les deux techniques suivent sensiblement la même méthodologie, celle-ci étant légèrement différente de celle développée dans notre étude. Elle comprend, en effet, une phase de validation de dimensions experts par les sujets, mais est aussi basée sur des séquences de verbalisation sans, puis avec utilisation/démonstration de sites portail. En reprenant le rapport de synthèse (Pican, 2001) et le questionnaire d'évaluation (construit suivant les étapes habituelles du processus et notamment l'exercice de hiérarchisation), nous observons qu'une assez grande partie des données utiles au questionnaire étaient déjà produites dans les *focus groups*. Les divers points traités dans ce questionnaire (la facilité à se repérer, l'organisation de la page d'accueil, l'esthétique, l'efficacité de la recherche à partir du site portail, l'efficacité des outils de communication comme le chat ou un forum, la publicité) sont évoqués par les participants des *focus groups*. Quelques détails concernant ces points et plusieurs adjectifs des réponses à choix multiples du questionnaire sont, en revanche, manquants.

Néanmoins, à notre sens (et malgré l'avis défavorable des professionnels de l'évaluation à l'égard des *focus groups*), ce questionnaire aurait pu, en grande partie, être élaboré à partir des données des *focus groups*. Ainsi, nous suggérons d'envisager la piste selon laquelle les données des *focus groups* permettraient d'obtenir directement les dimensions les plus importantes et surtout de construire le questionnaire d'évaluation. Cela serait envisageable uniquement en menant environ quatre à cinq *focus groups* avec des sujets de l'échantillon à étudier¹¹⁵. Il serait aussi nécessaire que les professionnels de l'évaluation aient une connaissance suffisante du domaine de produits/services et de leurs utilisateurs. Cette connaissance leur permettrait d'éviter de poursuivre de fausses pistes, entraînés par

d'éventuels participants exerçant un effet de leadership et de formuler les questions de façon plus appropriée.

Si nous suggérons cette piste, c'est aussi parce que l'intérêt de la situation de groupe réside dans le fait de produire, non pas une accumulation de points de vue individuels mais une construction collective (ex. Blanchet & Trognon, 1994 ; Maguire & Bevan, 2002). Comme nous l'avons vu (au chapitre 1), le *focus group* peut être utilisé comme outil de recherche de consensus pour produire, par exemple, une liste commune de descripteurs de produits alimentaires. De plus, le *focus group* peut révéler les valeurs partagées et priorités d'un groupe (Kitzinger, 1994) comme les accords et désaccords entre participants, et permettre de connaître le degré de consensus dans le groupe sur un thème donné (Morgan & Krueger, 1993). Nous avons aussi constaté (sans pour autant les relever systématiquement) des approbations générales ou le contraire qui permettraient de savoir si les regroupements opérés sont partagés par une majorité de participants ou non, voire de donner des indices sur l'importance relative de chaque regroupement.

Néanmoins, de plus amples recherches sont nécessaires pour vérifier l'hypothèse selon laquelle les méga-dimensions des *focus groups* pourraient correspondre aux dimensions les plus importantes et ainsi réduire le processus des professionnels de l'évaluation. Dans notre étude sur les sites Internet marchands, seule l'extraction des dimensions utilisateurs a été réalisée. La poursuite du processus, notamment l'exercice de hiérarchisation des dimensions et la construction du questionnaire n'ont pas eu lieu (pas de demande de production de résultats par des professionnels destinataires). Les dimensions les plus importantes concernant les sites Internet marchands n'ont donc pas été générées. Il n'est pas possible de comparer ces dernières avec les méga-dimensions produites en *focus group* ou en entretien individuel.

Cependant, il est possible de comparer les méga-dimensions produites en *focus group* et celles produites en entretien individuel. Elles sont différentes (cf. tableau 9, § 4.2.4.). Or, il faut éviter le piège d'utiliser les résultats de l'entretien individuel comme une référence pour évaluer le *focus group*. Il faut comprendre les différences obtenues (Kitzinger, 1994 ; Morgan, 1993). C'est ce que nous nous employons à faire dans ce qui suit.

7.3.2. Type de questionnaire produit à partir des méga-dimensions de chaque technique

Nous examinons ici quel type de questionnaire serait produit à partir de chacune des deux listes de méga-dimensions.

Les méga-dimensions obtenues à la suite des *focus groups* porte sur les principaux aspects relatifs à l'achat par Internet : les aspects commerciaux (les aspects sécurité et confidentialité, l'intérêt de l'offre de produits, des services après-vente, etc.), mais aussi des aspects émotionnels divers (crédibilité du site, esthétique, etc.) et de rares et généraux aspects relatifs à l'utilisabilité (rapidité de chargement, facilité de navigation). Ces diverses dimensions correspondent à celles des questionnaires généraux les plus fréquemment utilisés dans le domaine du Web marchand. Par exemple, le questionnaire SERVQUAL (Parasuraman *et al.*, 1985, 1991, 1993) comprend cinq dimensions qui cherchent à évaluer la perception de la fiabilité du service rendu par le site marchand (*Reliability*), de la capacité du site marchand à aider les clients et à leur fournir un service rapide (*Responsiveness*), du niveau de confiance accordé au site (*Assurance*), de la personnalisation du service rendu (*Empathy*) et enfin de tout ce qui relève de la conception graphique et fonctionnelle du site (*Tangibility*). Cela peut encore être conforté par la synthèse des dimensions les plus fréquemment employées dans plusieurs travaux du domaine : la confiance accordée au site marchand, la perception du risque, la taille perçue du site, la perception des produits mis en vente (prix, qualité, etc.), la qualité perçue du service et/ou du système, le plaisir, l'intention d'acheter, l'attitude face au site marchand ou face à l'informatique en général.

Deux questionnaires développés et utilisés par les professionnels du Sitemarchand.com, brièvement décrits dans les chapitres 5 et 6, relèvent aussi d'un ensemble de dimensions très générales, relatives aux divers aspects évoqués ci-dessus. En effet, nous découvrirons que les professionnels du Web marchand sont très concernés par l'ensemble des méga-dimensions issues des *focus groups* et que la question de l'utilisabilité constitue une seule de leur préoccupation parmi d'autres (la notoriété de leur commerce, l'appréciation du choix de produits qu'ils vendent, l'efficacité des services accompagnant l'achat, etc.).

Enfin, la situation collective permet de dégager très rapidement ces méga-dimensions. Une synthèse de la liste des titres formulés dans chaque *focus group* suffit. D'ailleurs, plus généralement, la productivité du *focus group* est meilleure que celle de l'agrégat d'entretiens individuels concernant la production de données d'un niveau général (niveau de la dimension).

Ainsi le *focus group* a son intérêt dans la construction d'un questionnaire, s'il s'agit d'une évaluation exploratoire générale, interrogeant la satisfaction des utilisateurs concernant divers aspects (évaluation multidimensionnelle à un niveau général), comme le font habituellement les professionnels du Web marchand. De plus, en sciences sociales appliquées et en marketing, le *focus group* présente un intérêt avéré pour la construction d'un questionnaire (Fuller *et al.*, 1993; Merton *et al.*, 1956; Morgan, 1997; O'Brien, 1993; Parasuraman *et al.*, 1985).

En revanche, les huit dimensions les plus partagées dans les entretiens individuels portent majoritairement sur l'utilisabilité et sont donc utiles dans le cas où une évaluation concernant exclusivement l'utilisabilité des sites marchands est envisagée.

Par ailleurs, ces méga-dimensions ne sont pas obtenues rapidement. Contrairement au *focus group*, les données obtenues lors d'un entretien individuel sont une accumulation de données sans connexion avec celles obtenues auprès d'autres sujets. Il a donc fallu identifier les dimensions qui étaient partagées par suffisamment de sujets et celles qui ne l'étaient pas, après un codage complet des données.

Enfin, toutes ces conclusions doivent être nuancées à partir des limites que présente cette étude.

7.4. LES LIMITES DE L'ETUDE

Cette étude présente certaines limites concernant la méthodologie du *focus group*. Elles sont explicitées dans ce qui suit. Nous abordons aussi la question du niveau de généralisation possible à partir de cette étude.

7.4.1. Concernant la méthodologie du focus group

La durée des *focus groups* était assez courte au regard de tous les exercices inclus. Il aurait fallu pouvoir laisser davantage de temps aux discussions sur les regroupements d'éléments individuels pour faire émerger les consensus par eux-mêmes. Faute de temps, les participants étaient parfois pressés de trouver des compromis, même si, in fine, les éléments qui suscitaient des désaccords étaient laissés à part (à côté d'un regroupement dont ils étaient susceptibles de faire partie mais sur lequel il n'y avait pas unanimité). De même, davantage de temps aurait été nécessaire aux discussions des participants observant le parcours réalisé

sur les sites marchands par un volontaire. Cette séquence présentait une certaine confusion avec les observateurs donnant des indications de parcours au surfeur et les discussions-commentaires sur les pages Internet émergeant un peu de toute part.

De plus, lors de la séquence d'utilisation/démonstration des deux sites marchands, un seul participant avait l'occasion d'utiliser le site marchand. Les autres observaient et cela n'est pas la même chose que s'ils avaient eu une réelle expérience d'utilisation (Nielsen, 1997). Cela explique, en partie, que davantage de méga-dimensions de nos *focus groups* portent sur des aspects généraux de l'achat par Internet et celles des entretiens individuels davantage sur l'utilisabilité.

Enfin, il est souvent préconisé de mener un minimum de trois *focus groups* (Krueger & Casey, 2000; Langford & McDonagh, 2003a; Urban & Hauser, 1993) pour pallier aux éventuelles défaillances de la situation de groupe (influence de meneurs de groupe, discussions hors sujet), même si des auteurs soutiennent qu'en ergonomie un *focus group* est suffisant (Caplan, 1990). Ainsi, la réalisation de plus de deux *focus groups* serait utile à la consolidation des résultats de cette étude. Pour une comparaison valide et réaliste, l'idée était de comparer l'une et l'autre technique avec la même quantité de sujets (de même que Fern, 1982, et que les professionnels de l'évaluation), mais cela amène, en réalité, à comparer un volume verbal bien moindre en *focus group* qu'en entretien individuel.

Ces lacunes ont pu, en partie, jouer sur les performances précédemment observées, même si nous avons considéré les proportions d'idées nouvelles, détaillées, etc. par rapport à la quantité totale d'idées de *chaque* technique. D'autres résultats (même s'ils ne tiennent pas compte de la productivité des techniques) peuvent néanmoins faire réfléchir. Par exemple, il s'avère que davantage de symboles pertinents sont créés en *focus group* plutôt qu'avec une méthode individuelle, pour la conception de symboles picturaux pour des applications logicielles pour des techniciens de maintenance dans l'aviation (Macbeth *et al.*, 2000).

7.4.2. Niveau de généralisation possible à partir de l'étude

Mise à part les précédentes réserves, il s'agit de rester vigilant quant à la généralisation des résultats de cette étude. Les conclusions de la supériorité de l'entretien individuel par rapport au *focus group* pour la production d'un maximum d'idées détaillées (ou vice versa, concernant la production de données d'un niveau général) sont circonscrites au cadre d'une étude qualitative où il s'agit de générer des idées et pensées. Cette étude est destinée à

élaborer une liste de dimensions utilisateurs et des connaissances sur les internautes et cyber-consommateurs (qui peuvent aussi servir à renseigner des besoins utilisateurs¹¹⁶). Des tâches de génération d'idées ou impressions à partir d'expériences du quotidien ou lors de l'utilisation du produit/service sont données aux sujets. Les conclusions développées ici ne sont a priori pas généralisables ou « transférables » (Lincoln & Guba, 1985) si les sujets sont soumis à des tâches très différentes de celles de l'étude.

De plus, à première vue, il semble acceptable de généraliser les conclusions de cette étude à toute autre démarche d'élicitation d'idées (voire de besoins utilisateurs) pour tout produit/système ou tout produit/service. Cependant, cela mériterait de plus amples recherches, étant donné que les effets des techniques d'élicitation sur les attributs générés peuvent être différents selon qu'il s'agisse de services (ex. restaurant) ou de produits (ex. voiture), par exemple, en termes de variété des attributs et de leur « pouvoir » de discrimination (Breivik & Supphellen, 2003).

Après ces remarques de prudence, nous élargissons l'interprétation de certains de nos résultats à l'utilisation de ces techniques pour la conception de produits/services.

8. INTERET DES DEUX TECHNIQUES POUR LA CONCEPTION DE PRODUITS/SERVICES

Malgré les précédentes remarques, « on limiterait la portée de la phase qualitative si on n'en faisait qu'une étape de préparation. Malgré le petit nombre de personnes interrogées, certaines conclusions suffisamment solides peuvent être tirées... » (Ghiglione & Matalon, 1998, p. 93). En plus des données utiles à la construction du questionnaire, nous obtenons, en effet, des données qui sont de l'ordre de résultats d'évaluation des sites marchands et présentent un intérêt direct pour les professionnels du marketing ou de la conception.

Nous élargissons donc nos interprétations et débouchons notamment sur l'intérêt, d'une part, de combiner les deux méthodes et, d'autre part, de favoriser les discussions de confrontation en *focus group*. Certaines de ces conclusions restent valables dans le champ de la construction de questionnaires.

8.1. COMBINER DES ENTRETIENS AVEC UNE SEQUENCE D'UTILISATION INDIVIDUELLE ET DES *FOCUS GROUPS*

Nous mettons d'abord en évidence les avantages des données obtenues dans chaque technique pour être directement utilisées dans la conception. Nous observons ici plus particulièrement les données obtenues lors des séquences d'utilisation/démonstration des sites marchands, puis afin de conjuguer ces avantages, nous proposons une solution de combinaison des deux techniques.

8.1.1. Les données pour la conception issues des entretiens individuels

En entretien individuel, la phase d'utilisation apporte légèrement plus d'idées détaillées que ne le fait la première phase sans utilisation du produit/service. Ces données, issues de l'analyse des verbalisations en référence à l'activité, sont riches en éléments d'utilisabilité, même s'il ne s'agit pas de mesures de performance.

Ces données sont donc directement utiles pour la conception de l'interface. Les concepteurs ou chefs de projet aiment prendre des décisions basées sur des données quantitatives pour allouer des ressources financières là où il y aura le plus de valeur ajoutée (Karat, 1994), et les tests utilisateurs peuvent fournir ces éléments. Cependant, à partir de la seule analyse de l'activité lors de nos entretiens individuels, il y a moyen d'indiquer la gravité d'un problème d'utilisabilité (engendrant seulement un ralentissement dans l'accomplissement de la tâche ou un échec complet) d'une classe de situation donnée (ex. sélection des caractéristiques du produit à mettre dans le panier de commande) et de préciser si tel ou tel problème a été rencontré par une majorité d'utilisateurs ou non.

Ces entretiens individuels peuvent donc avoir un double intérêt : produire le matériel nécessaire à la construction du questionnaire et produire certaines données directement utiles aux acteurs de la transformation.

8.1.2. Les données pour la conception issues des focus groups

Concernant les *focus groups*, l'ensemble des données obtenues permet de déterminer quels sont les grands points sur lesquels il s'agit de travailler (cf. méga-dimensions) pour améliorer les sites commerciaux sans pour autant fournir de détails précisant comment il faut

les concevoir. En effet, le *focus group*, au contraire des entretiens ou d'autres techniques, est connu pour ne pas fournir de données détaillées pour diriger la conception d'un produit (Van Vianen *et al.*, 1996) ou du moins pour permettre d'orienter les choix de conception sans fournir de données fiables sur l'utilisabilité d'un produit (Baccino, *et al.*, 2005; Nielsen, 1997). Pour concevoir et/ou évaluer correctement un produit/service, les *focus groups* doivent généralement être complétés par des données quantitatives d'utilisabilité (ex. Nielsen, 1999).

Par ailleurs, la séquence d'utilisation/démonstration des *focus groups* a produit bien davantage de données nouvelles que la première phase d'entretien. Or, les deux phases, avec et sans utilisation, présentent quelques défaillances. Nous ne pensons pas que la première phase en présente davantage que la seconde, mais plutôt que les exercices de catégorisation d'éléments individuels et de partage d'expériences vécues ne sont pas aussi concrets que les échanges à partir de l'observation d'un parcours réel sur un site marchand. Ainsi, la séquence d'utilisation/démonstration aurait permis aux participants d'être plus « innovants ».

Pour conserver cet effet positif sans en avoir les inconvénients (confusion, un seul participant qui navigue sur les sites), une solution consisterait à combiner des entretiens avec une séquence d'utilisation individuelle à partir d'un scénario de tâche, puis des discussions de groupe à partir de cette même expérience. Il s'agit d'ailleurs d'une solution largement mise à contribution dans la conception de produits/services (Bouch *et al.*, 2000; Caplan, 1990; Coughlan & Sklar, 2003; Jarvenpaa & Todd, 1997; Langford & McDonagh, 2003a, 2003b; O'Donnell *et al.*, 1991)¹¹⁷. Cela permettrait également de pallier à un éventuel écart entre ce que les gens disent vouloir et ce dont ils ont réellement besoin (ex. Nielsen, 1993).

Cette solution de combinaison des deux techniques peut être pertinente aussi bien pour produire des données utiles à la conception que pour construire un questionnaire d'évaluation. En effet, les deux techniques peuvent être avantageusement combinées pour préparer une enquête et recueillir les données ou même pour vérifier les données produites (Kitzinger, 1994; Morgan, 1997).

Nous présentons maintenant un autre point fort des *focus groups* : les discussions de confrontation.

8.2. LES DISCUSSIONS DE CONFRONTATION POUR ETUDIER UNE VARIETE D'USAGES ET DE MOTIVATIONS

Nous pensons que la part d'explications spécifiques aux *focus groups*, bien que faible, n'aurait pas été élaborée en entretien individuel (Kitzinger, 1994; Krueger & Casey, 2000; Langford & McDonagh, 2003a). Ces éléments ont été exprimés majoritairement lors de confrontations entre les participants concernant leurs habitudes et/ou motivations dans l'utilisation des sites marchands.

Il nous semble donc intéressant de creuser cette piste des discussions de confrontation. Il s'agirait de favoriser de telles discussions (Mariampolski, 1991) pour traiter certains thèmes cruciaux requérant une confrontation détendue de points de vue relatifs à différents usages et motivations.

Il s'agirait de rassembler, par exemple, des internautes sensiblement de la même catégorie socioprofessionnelle, ayant la même facilité et volonté de s'exprimer sur le thème donné, mais dont on sait qu'ils ne partagent pas les mêmes usages (ex. ceux qui préfèrent utiliser une recherche par mots-clefs, ceux qui préfèrent utiliser les liens) ou de différentes classes d'âge (ce qui favorise a priori les usages ou motivations divergentes). Ces groupes en partie hétérogènes doivent être construits avec attention, de façon à ne pas provoquer de blocage dans les interactions (ex. personnes très intimidées face à des personnes de classe sociale plus élevée)¹¹⁸.

De plus, il s'agirait de mener des *focus groups* en ciblant quelques thèmes importants seulement, ainsi que le recommandent Cooper et Baber (2003). Ces thèmes peuvent être identifiés lors d'entretiens et/ou d'observations préalables de l'activité. Dans le domaine des sites marchands, ils peuvent, par exemple, porter sur des sujets sensibles comme les préférences en termes d'options de sécurité proposées (ex. enregistrement ou non du numéro de carte bancaire dans les fichiers de l'entreprise marchande¹¹⁹) ou en termes de fonctionnement de la recherche (par mots-clefs, par orthographe voisine ou non, etc.).

Il est généralement recommandé de mener des groupes homogènes (ex. Caplan, 1990; Maguire, 2003). Cependant, Maguire (2003) observe qu'un participant plus jeune et plus expérimenté dans les technologies de l'information par rapport aux autres participants d'un *focus group* contribue à faire émerger des discussions intéressantes car il est moins sceptique que ces collègues concernant ces technologies. Dans le domaine des sciences sociales, de plus amples recherches seraient utiles pour savoir comment l'homogénéité ou l'hétérogénéité des

groupes affectent ce que les participants sont susceptibles de dire (Morgan, 1993). Enfin, Mariampolski (lors d'échanges informels) confirme que ce point, soulevé dans les années 1990, n'a été depuis, à tort selon lui, approfondi ni par lui-même ni par autrui.

Cela constitue, de notre point de vue, un axe de recherche à explorer davantage. Nous suggérons de comparer des *focus groups* parfaitement homogènes à des *focus groups* présentant une certaine hétérogénéité et d'examiner si ces derniers produisent davantage d'idées pertinentes que les autres *focus groups*.

9. CONCLUSION

Tout d'abord, cette étude a permis de construire une nouvelle liste détaillée de dimensions de perception utilisateurs¹²⁰, destinée aux professionnels de l'évaluation. Si une demande d'évaluation concernant des sites marchands leur est soumise, ils peuvent s'en servir pour créer un questionnaire d'évaluation (la formulation précise de chaque question devra faire l'objet d'un recueil plus systématique des termes les plus employés par les internautes). L'avantage de la liste exhaustive détaillée (trois niveaux de profondeur) est de leur permettre de créer aussi bien un questionnaire global que plus spécifique, selon les besoins, même si une évaluation exploratoire générale leur est plus souvent demandée.

Une conclusion plus générale est que l'entretien individuel est préférable au *focus group*, pour l'élaboration d'une liste exhaustive de dimensions de perception avec une grande précision. De plus, la productivité de l'entretien individuel en termes de coûts financiers et humains est meilleure que celle du *focus group*. Il semble recommandé d'utiliser ces entretiens individuels dans la phase qualitative amont à la construction d'un questionnaire, en particulier s'il s'agit d'une évaluation ciblée sur l'utilisabilité de produits/systèmes et si les évaluateurs ont une faible connaissance du domaine (les entretiens individuels produisent de nombreuses données détaillées). Enfin, il s'agirait de consolider ces résultats avec d'autres et notamment de comparer des quantités plus équitables de données produites par l'une et l'autre technique. Néanmoins, nous pouvons affirmer que l'entretien individuel et le *focus group* peuvent être avantageusement combinés.

En revanche, nous suggérons aussi aux professionnels du service d'évaluation (cela les concerne plus directement) de creuser encore la piste de l'utilisation des *focus groups*. Ces derniers, très productifs concernant l'élaboration de données d'un niveau général, peuvent peut-être produire une liste synthétique de dimensions directement utile à la construction du

questionnaire. Cela, uniquement dans le cas où une évaluation exploratoire générale est requise, à savoir qui porte sur une variété d'aspects généraux d'un produit/service (ex. la notoriété/crédibilité d'un site marchand, la sécurité, l'intérêt des produits/services après-vente, l'esthétique du site, la rapidité de chargement, etc.). Cela permettrait aux professionnels de faire l'économie de la phase de production d'une liste exhaustive de dimensions qui est ensuite réduite. Cela pourrait être possible à condition de mener suffisamment de *focus groups* (quatre à cinq) avec uniquement les sujets concernés par l'étude finale et non différents groupes d'utilisateurs pour recueillir la plus grande diversité d'opinions et de dimensions possibles. Suite à l'étude, c'est encore une simple hypothèse qui nécessite d'être vérifiée.

De plus, nous avons démontré qu'il est pertinent de combiner des entretiens avec une séquence d'utilisation individuelle et des discussions de groupe pour produire des résultats pour la conception. Enfin, un axe de recherche, encore peu exploré et présentant un certain intérêt, sont les *focus groups* en partie hétérogènes (mais soigneusement construits) dont les discussions de confrontation semblent générer des idées non obtenues en entretien individuel.

Dans cette étude, nous n'aborderons qu'un seul versant de l'évaluation, à savoir la production de dimensions et (par extrapolation) de résultats d'évaluation. Dans la suite de cette thèse, nous examinons la question de leur utilisation et, plus concrètement, la nécessaire opérationnalisation des dimensions utilisateurs pour que des professionnels de la transformation (notamment du marketing) puissent les utiliser pour concevoir leurs propres résultats d'évaluation.

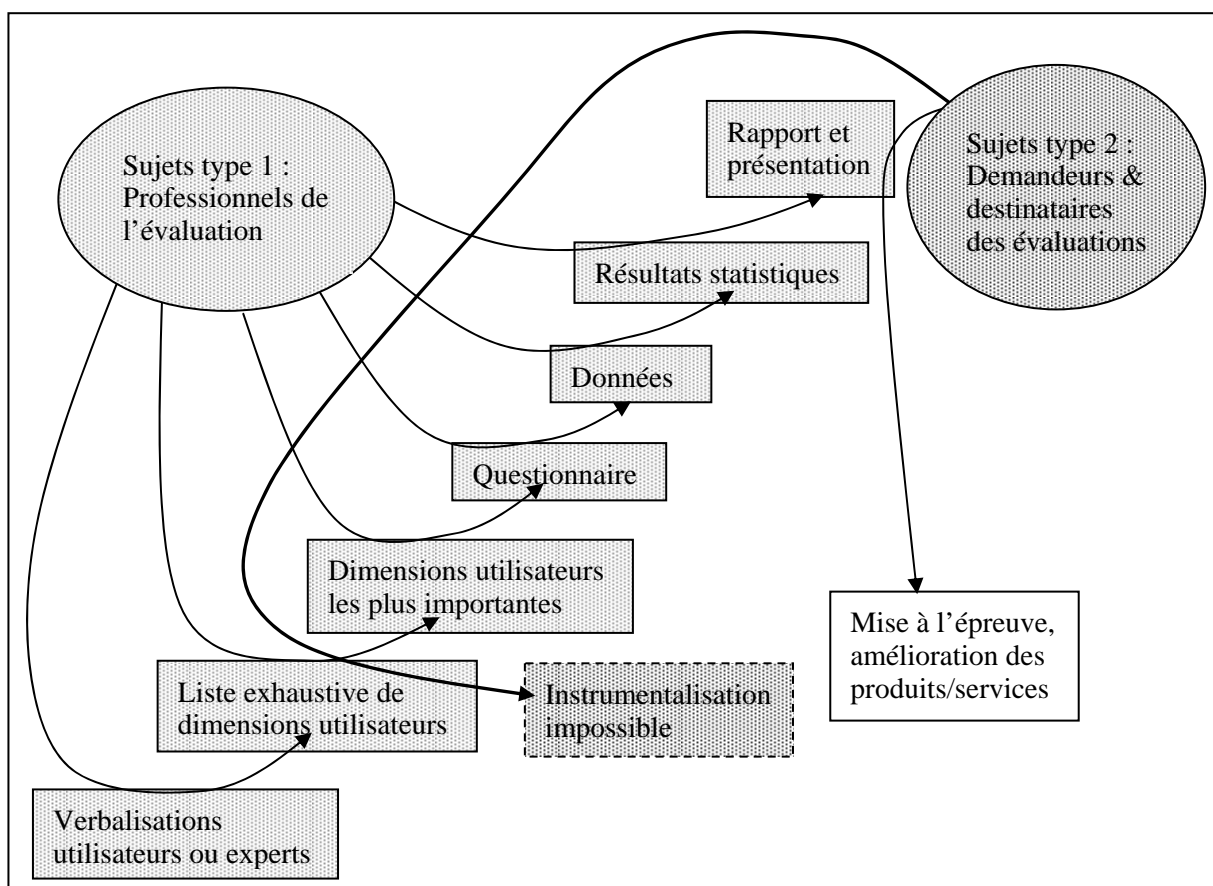
CHAPITRE 4

DEUX POINTS DE VUE SUR L'ÉVALUATION

REPERES POUR LE CHAPITRE

Nous réutilisons le schéma modélisant le processus d'évaluation depuis l'extraction des dimensions utilisateurs jusqu'à la restitution du rapport aux professionnels destinataires. Les parties grisées de la figure 14 (ci-dessous) représentent les parties du processus concernées par ce chapitre. Nous exposons, d'une part, l'activité des professionnels de l'évaluation (sujets de type 1) et leur processus d'évaluation en entier (parties en gris clair). D'autre part, nous montrons que les professionnels de la conception et du marketing (destinataires potentiels et donc sujets de type 2) ne parviennent pas à utiliser, donc à instrumentaliser les dimensions utilisateurs (parties en gris foncé).

Figure 14 : Processus élaboré par les professionnels du service d'évaluation, depuis l'extraction des dimensions utilisateurs jusqu'à la restitution des résultats d'évaluation aux professionnels destinataires



Légende : Les sujets de type 1 ont une activité d'évaluation et les sujets de type 2 ont une activité de conception/marketing des produits/systèmes.

1. INTRODUCTION

Après avoir généré une liste exhaustive de dimensions utilisateurs et discuté de la performance de l'entretien individuel et du *focus group*, nous abordons le renversement de point de vue porté sur l'évaluation. Nous nous intéressons à l'activité des professionnels du service d'évaluation, d'une part, et à celle de professionnels engagés dans le marketing et la conception de sites Internet, d'autre part. Ces derniers, en particulier les professionnels de la veille-marketing, ont leur propre activité d'évaluation puisqu'ils évaluent eux-mêmes des sites Internet sans avoir recours aux professionnels de l'évaluation.

L'objet de ce chapitre est de mettre en évidence deux points de vue concernant la question de l'évaluation.

Premièrement, pour les professionnels de l'évaluation, isolés de l'activité de conception et de marketing des produits/services, la liste de dimensions utilisateurs constitue le matériau de fabrication de l'évaluation. Ces dimensions ou critères de jugement des utilisateurs/clients correspondent à ce qu'il faut évaluer. La production et l'usage de la liste de dimensions utilisateurs s'inscrivent dans le processus d'évaluation, en partie présenté au chapitre précédent. Ce processus est pensé par ces mêmes acteurs en lien avec un contexte économique et organisationnel plus large, relevant d'un certain cloisonnement entre les entités demandant une évaluation, réalisant l'évaluation et assurant la transformation des produits/services. Tel un procès de distillation, le processus d'évaluation doit, en tous cas, déboucher sur des résultats d'évaluation considérés comme « substances » essentielles et remis à divers professionnels de la transformation des produits/services qui s'en serviront. L'ensemble de cette démarche est centré exclusivement sur un artefact, la liste de dimensions utilisateurs, sans considérer la réalité de l'activité des professionnels destinataires. Il y a production de connaissances sans penser en amont l'utilisation de ces connaissances.

Deuxièmement, les professionnels du marketing, ancrés dans l'activité de transformation de sites marchands, se basent sur ce qu'ils savent concrètement avoir besoin d'évaluer dans leur propre activité d'évaluation à partir des connaissances de leurs utilisateurs/clients et des contraintes inhérentes à leurs produits/services. Leur souci est celui de l'efficacité et de l'utilité : la production de connaissances recueillies auprès des utilisateurs/clients est pour eux subordonnée à l'agir. Ils ont besoin que soient produits les résultats d'évaluation dont ils peuvent effectivement se servir.

Dans ce chapitre, nous présentons d'abord brièvement les professionnels que nous avons côtoyés ou interviewés. Nous décrivons ensuite le processus d'évaluation dans son ensemble, ainsi que le point de vue des professionnels de l'évaluation et les facteurs économiques et organisationnels participant à ce point de vue. Ce dernier est aussi discuté au regard de la question de la « vérité » et de l'utilité d'un résultat d'évaluation. Dans un second temps, une conduite itérative d'entretiens avec les professionnels de la veille-marketing et de la conception de sites Internet est détaillée, ainsi que la transformation de surface opérée sur la précédente liste exhaustive de dimensions utilisateurs des sites Internet marchands¹²¹ que nous leur avons présentée. Nous démontrons, en analysant de plus près les données recueillies, que chaque transformation opérée est insuffisante pour les professionnels de la veille-marketing et qu'ils ont, en fait, besoin d'un outil opérationnel de conception d'une évaluation. Nous développons aussi l'hypothèse de deux mouvements épistémologiques distincts à l'œuvre dans les représentations et activités des professionnels de l'évaluation et du marketing, en lien avec les deux notions de la vérité et de l'utilité d'une évaluation.

2. PRESENTATION DE LA DEMARCHE SUIVIE

2.1. UNE DEMARCHE « NON METHODOLOGIQUE »

En premier lieu, notre constat se base sur nos trois ans de travail au sein du service d'évaluation de la branche FTA. Nous nous appuyons sur de nombreuses discussions informelles et sur quelques précédents travaux menés avec les professionnels de l'évaluation. Nous sommes en mesure d'avancer certains éléments les concernant sans pouvoir toutefois exposer de données précises (verbalisations retranscrites ou autres).

Ce travail se base ensuite sur une série d'entretiens menés avec deux professionnels de la veille-marketing et trois professionnels de la conception de sites Internet (marchands) pour recueillir leurs réactions face à l'artefact précédemment élaboré, la liste de dimensions utilisateurs. La démarche suivie concernant ces entretiens est très empirique, les premiers entretiens nous amenant à en poursuivre d'autres pour approfondir des éléments lacunaires. C'est ce que nous appelons une conduite itérative d'entretiens. Dans un premier temps, à partir d'une première analyse sommaire des entretiens (retranscription partielle des entretiens), nous avons procédé à une transformation de surface de la liste de dimensions utilisateurs cherchant simplement à ce que cette liste puisse être appréhendée plus aisément par les acteurs de la veille-marketing et de la conception. C'est au cours de ce processus de

conception que certaines remarques des professionnels, de la veille-marketing notamment, nous font penser que cet aménagement est insuffisant. Nous avons alors retravaillé à partir des données entièrement retranscrites et certains éléments nous conduisent à constater que cette liste (transformée ou non) ne correspond pas à leurs besoins en termes d'outils. Ils ne peuvent pas s'en servir telle qu'elle. Elle n'est pas instrumentalisable par eux. Nous en tirons la conclusion qu'il est nécessaire de développer un outil opérationnel pour les professionnels du marketing (et de la conception).

Nous présentons quelques données issues de ces entretiens dans ce chapitre et de nombreuses autres, utiles à la construction de cet outil, dans les deux suivants.

2.2. PRESENTATION DES PROFESSIONNELS SOLLICITES

2.2.1. Les professionnels côtoyés sur le premier terrain de recherche : les professionnels de l'évaluation

Comme nous l'avons précisé, ces professionnels du service d'évaluation de la branche FTA ont exclusivement une activité d'évaluation des produits/services conçus dans d'autres services ou branches opérationnelles et sont isolés de l'activité de transformation. Ils ont essentiellement des interactions avec les demandeurs des études lors de la négociation et spécification des études, et avec les acteurs de la transformation lors de la restitution des résultats et conclusions d'études.

2.2.2. Les professionnels interviewés sur le second terrain de recherche : les professionnels de la transformation

Certains de ces professionnels de la transformation sont aussi des acteurs de l'évaluation.

Le premier professionnel du service de veille-marketing, spécialiste du Sitemarchand.com

Le premier professionnel est le spécialiste de la veille-marketing du Sitemarchand.com, il assure une activité d'études stratégiques (enquête par questionnaire, mesures d'audience sur le site) et une activité de conseil auprès des professionnels des équipes marketing et Webmastering de l'entreprise du Sitemarchand.com. Il restitue notamment les données d'enquêtes ou diverses connaissances concernant des sites Internet ou nouvelles technologies

de France Télécom. Il participe également à l'activité d'évaluation assurée par l'entreprise Sitemarchand.com de son propre site. Il est un acteur de l'évaluation.

Le second professionnel du service de veille-marketing, non spécialiste du Sitemarchand.com

Le second professionnel du service de veille-marketing n'a pas les connaissances du précédent concernant l'activité « bout en bout » d'une entreprise marchande par Internet. Il assure néanmoins une activité d'études de marché et de veille-marketing d'orientation pour divers projets Internet, y compris concernant des activités commerciales par Internet. Il a aussi une activité de régie de la publicité en ligne (« achat » et vente de panneaux publicitaires sur le Web) pour la filiale qui est conceptrice de ses propres publicités.

Le responsable du Webmastering du Sitemarchand.com

Comme déjà précisé, l'activité du responsable du service Webmastering de l'entreprise Sitemarchand.com concerne la spécification de la conception du Sitemarchand.com. Par ailleurs, le responsable du Webmastering, peut commander des études d'évaluation concernant spécifiquement le site Internet (interface, recherche par mots-clefs, etc.) et contribuer à spécifier les « besoins en évaluation » relatifs à la conception lors des régulières évaluations par questionnaire. Il n'est pas pour autant un acteur de l'évaluation au même titre que le professionnel de la veille-marketing ou que le chargé des études du Sitemarchand.com (cf. chapitre 2).

Deux ingénieurs concepteurs de sites Internet

Les deux ingénieurs sont « à part » quant au découpage des deux terrains de recherche. Ils exercent leur activité au sein de services de la branche FTA, indépendants du service d'évaluation, et nous avons aussi simplement mené des entretiens avec eux. L'un des ingénieurs a conçu l'interface d'un site Internet portail de France Télécom (son architecture et chacune de ses pages HTML¹²²) et connaît bien le nécessaire rôle d'intermédiaire entre le marketing, les infographistes et les développeurs du « back-office » (programmes informatiques en arrière-plan) inhérent à cette fonction. L'autre est davantage spécialisé dans les études de faisabilité dans le domaine du Web marchand (ex. plate-forme sécurisée pour les comptes d'achat créés par les internautes) et dans le développement informatique du « back-office ».

Nous abordons d'abord le premier point de vue sur l'évaluation, porté par les professionnels de l'évaluation.

3. LA CONSTRUCTION D'UNE EVALUATION A PARTIR DES DIMENSIONS UTILISATEURS

3.1. L'ARTEFACT « DIMENSIONS UTILISATEURS » DANS CE PROCESSUS DE CONSTRUCTION

Nous décrivons en entier le processus d'évaluation développé par les professionnels du service d'évaluation à partir de notre connaissance de ce terrain. Pour cette description, nous utilisons la notion de médiation pragmatique empruntée à la théorie instrumentale (Rabardel, 1995), étant donné qu'il s'agit d'un processus de développement d'artefacts à l'aide d'une succession de transformations d'un artefact de départ, la liste de dimensions utilisateurs. Cette description concerne uniquement les évaluations exploratoires par questionnaire puisque nous avons pu observer ce processus précisément et même participer à deux études de ce type (une étude sur les sites Internet portail - Pican, 2001 - et étude d'identification des dimensions utilisateurs des sites Internet marchands - Clouet & Lambert, 2002).

3.1.1. Un contexte donné

Parmi toutes les études commandées au service d'évaluation, les évaluations exploratoires en phase de « feed-back » proviennent de demandes assez larges. Elles portent non pas sur un produit/service précis à évaluer à un temps « t » donné comme c'est le cas des professionnels du Sitemarchand.com, mais sur les habitudes d'usage et la perception d'un ou de plusieurs groupes d'utilisateurs/clients concernant des produits/services d'une même famille (ex. les sites Internet portail, les sites Internet marchands, etc.). La demande d'évaluation et les dimensions utiles à évaluer sont peu spécifiées. Les professionnels de l'évaluation se basent donc sur le principe d'une liste la plus exhaustive possible de dimensions utilisateurs et leur attribuent la propriété d'être les éléments pertinents à évaluer.

Hormis le contexte donné par la demande peu précisée, cela part du point de vue, clairement exprimé par les professionnels de l'évaluation, selon lequel il s'agit d'extraire les dimensions sur lesquelles reposent les opinions et motivations des utilisateurs des « tréfonds de leur cerveau » et de réaliser une évaluation suivant ces dernières pour parvenir à des

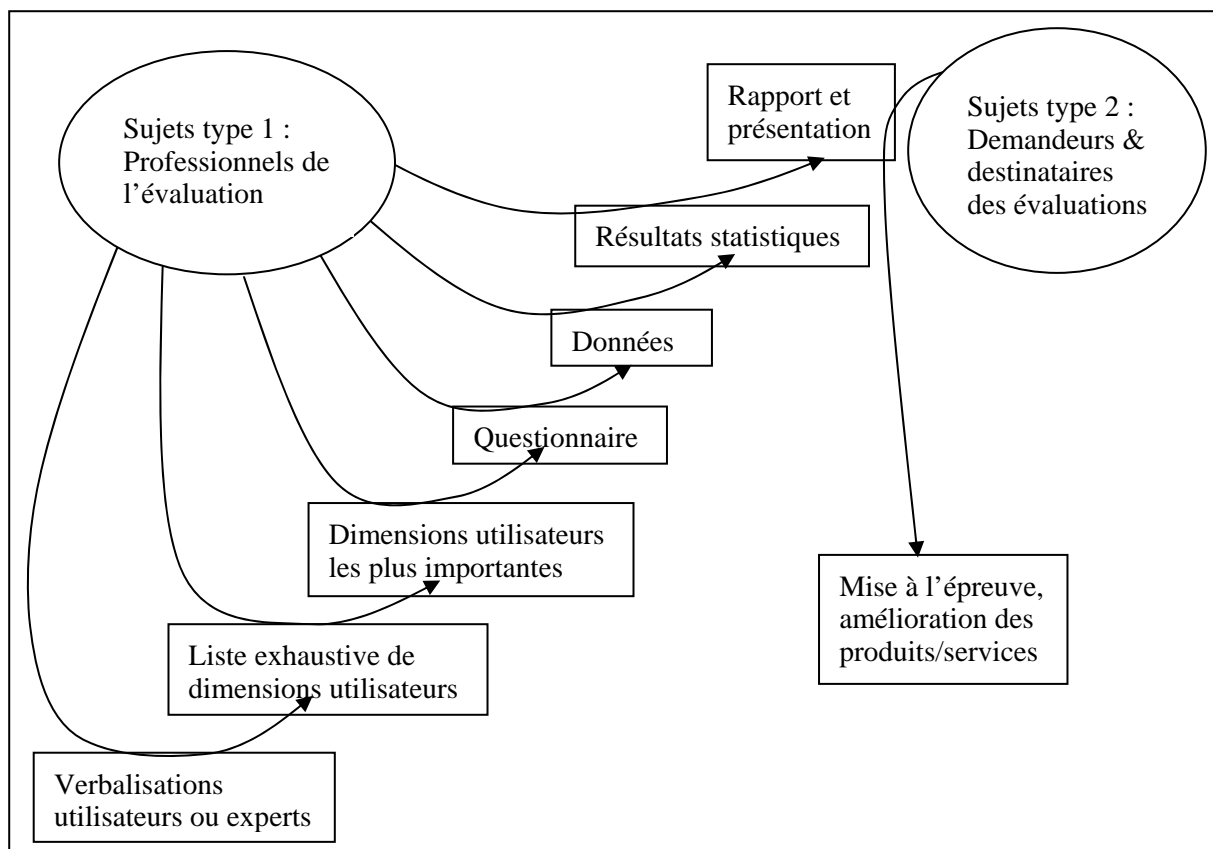
résultats qui reflètent au plus près les jugements des répondants. Par ailleurs, ces dimensions sont susceptibles d'être utiles sur le long terme, réutilisées pour d'éventuelles évaluations ultérieures. Dans une certaine mesure, elles ont un caractère figé.

3.1.2. Une série de médiations pragmatiques

Dans cette perspective, cette liste exhaustive de dimensions utilisateurs constitue la matière première permettant de générer le questionnaire d'évaluation. Cet artefact est le point de départ, le germe du processus. La liste exhaustive de dimensions utilisateurs est d'abord extraite à partir, d'une part, des dimensions élaborées par des experts du domaine de produits/services et, d'autre part, à partir des verbalisations recueillies auprès de plusieurs groupes d'utilisateurs. Un processus de réduction est ensuite opéré en faisant hiérarchiser les dimensions selon leur degré d'importance par un échantillon du seul groupe d'utilisateurs à étudier spécifiquement dans l'évaluation finale. Le questionnaire d'évaluation est construit à partir des seules dimensions les plus importantes qui sont traduites sous la forme de questions et agencées selon une série de règles de construction issues d'un savoir déjà formalisé dans de nombreux manuels d'enquêtes (cf. chapitre 1) ainsi que par les professionnels de l'évaluation en tenant compte de leur propre contexte d'application (Le Calonnec, 2001; Marchadour, 2001). Grâce à ce questionnaire, généralement administré par Internet, les réponses des sujets sont recueillies. Ces données sont dépouillées, compilées, parfois combinées, puis transformées en résultats statistiques selon les objectifs de l'étude et grâce à une série d'outils statistiques. Enfin, des règles de rédaction correspondant aux format et formulation attendus par les destinataires permettent la réalisation du rapport et de la présentation orale. Ces derniers sont, en effet, travaillés de façon à être le plus aisément appréhendés par ces destinataires, au niveau de son format (le plus concis possible) et dans sa formulation globale (éviter un inventaire de points uniquement négatifs, sans compromettre la réalité mise au jour par les résultats).

Il s'agit là d'un processus d'instrumentalisation à partir de la liste de dimensions, décomposable en transformations opérant chacune son action sur le produit de la précédente. Cette série de médiations pragmatiques aboutit d'abord à la liste exhaustive de dimensions, aux dimensions les plus importantes, au questionnaire, aux données, aux résultats statistiques et enfin au rapport et/ou présentation orale (cf. figure 15 ci-dessous).

Figure 15 : Processus élaboré par les professionnels du service d'évaluation pour les évaluations par questionnaire : série de médiations pragmatiques opérant à partir des dimensions utilisateurs



Légende : les flèches modélisent les médiations pragmatiques, les ronds les sujets et les carrés les artefacts issus des transformations ; les sujets type 1 ont une activité d'évaluation ; les sujets type 2 ont une activité de conception /marketing des produits/systèmes

Le processus d'évaluation est donc centré sur les dimensions utilisateurs extraites de la synthèse des connaissances des experts et des verbalisations des utilisateurs. Il s'agit de leurs jugements, préférences, exigences, etc. Ces dimensions correspondent donc aux catégories de perception des utilisateurs/clients et sont le vecteur de tout le processus puisqu'il s'agit d'évaluer selon ces dimensions utilisateurs. A la fin seulement et de façon superficielle, les destinataires sont pris en compte en appliquant un « filtre » correspondant à leur compréhension. La pertinence d'une évaluation, selon les professionnels de l'évaluation est donc considérée du seul point de vue des utilisateurs. Il n'y a pas d'analyse approfondie et de prise en compte en amont des réels besoins en connaissances des divers professionnels qui vont tenter d'utiliser les résultats et faire évoluer le produit/service.

3.2. DES FACTEURS ECONOMIQUES ET ORGANISATIONNELS

Nous pensons que cet état de fait est lié à plusieurs facteurs, sans pouvoir discerner précisément le poids relatif de chacun :

- des facteurs économiques,
- des facteurs organisationnels,
- et le point de vue adopté par les professionnels de l'évaluation.

Des facteurs économiques et organisationnels sont donc imbriqués dans le point de vue propre à ces professionnels, qui les amène à ne pas traiter la question de l'utilité des résultats produits.

3.2.1. Les facteurs économiques

Certains professionnels de l'évaluation, malgré leur parti pris, reconnaissent eux-mêmes leurs lacunes en matière de connaissances sur l'utilisation des résultats d'évaluation par les professionnels destinataires. Lors d'une réunion annuelle de bilan du service, ce point a fait l'objet d'une discussion entre plusieurs acteurs. Le fait est qu'ils ne disposent pas des moyens nécessaires pour analyser de plus près cette activité d'utilisation des résultats. Etant donné leur charge de travail, une unité dédiée à part entière à l'étude de ce versant du problème est indispensable. Pour des raisons évidentes de restriction budgétaire, la création d'une unité supplémentaire n'était cependant pas à l'ordre du jour.

3.2.2. Les facteurs organisationnels

De plus, étant donné le mouvement de privatisation de France Télécom, les rapports entre différents services, branches ou filiales sont de plus en plus des rapports commerçants. Ce phénomène n'est pas isolé. De nombreuses grandes entreprises connaissent cette logique, où chaque service est cloisonné, a ses propres responsabilités, budgets, frontières et comptabilités. Cela encourage peu le partage de connaissances et les études transversales au sein des entreprises (Lindgaard, 2002).

Nous faisons aussi l'hypothèse que ce cloisonnement entre les différentes entités opérant dans le processus « demande d'évaluation – réalisation de l'évaluation – conception » amène à ce que les besoins en connaissances des divers professionnels transformant les produits/services soient peu étudiés. Cela peut contribuer à renforcer le principe des

demandes d'évaluations exploratoires peu spécifiées, engageant les professionnels de l'évaluation à se baser sur les dimensions utilisateurs.

Pour évoquer le peu d'analyses des besoins en évaluation des professionnels, nous nous appuyons sur des éléments recueillis à partir :

- de deux entretiens retranscrits, menés par deux professionnels de l'évaluation, avec deux chefs de produit, dont un faisant partie d'une branche opérationnelle assurant la production, la maintenance et la diffusion d'un produit de France Télécom,
- d'un entretien mené par le chercheur avec un ingénieur de FTA ayant participé au développement informatique d'une application logicielle pour l'accès à Internet.

Une étude d'évaluation sur la satisfaction des utilisateurs quant à ce produit, a été commandée par la direction commerciale d'une branche opérationnelle de France Télécom aux professionnels de l'évaluation. Le rapport de synthèse a été remis à la direction commerciale de la branche, puis transmis aux concepteurs de FTA. L'entretien que nous avons conduit avec l'un des concepteurs commentant ce dernier rapport a mis en évidence que si plusieurs résultats confirmaient la pertinence des choix de conception, quelques autres n'étaient pas assez approfondis pour que les concepteurs puissent prendre des décisions en termes de modification du processus d'installation ou de l'aide intégrée (ex. les verbatim qui énoncent des problèmes sans données de hiérarchisation de l'importance relative des différents problèmes ou le pourcentage d'utilisateurs satisfaits sans détail permettant de savoir à quoi attribuer le résultat¹²³).

D'après l'entretien avec le chef de produit, avant la conception d'une nouvelle version du produit, ce dernier réalise un audit de personnes de la direction marketing et des autres directions de sa branche opérationnelle ayant utilisé ou entendu parlé de problèmes relatifs à la dernière version du produit. Il effectue la synthèse de divers éléments provenant d'un groupe analysant les remontées du service clients et de l'assistance technique, d'ingénieurs de la branche testant et améliorant le système technique, et d'éventuelles personnes regardant le fonctionnement général des produits/services de la concurrence. Il cherche à connaître les futures opérations commerciales à couvrir (ex. décor de Noël), les nouveaux besoins en conception (ex. fonctionnalités supplémentaires) et enfin les problèmes techniques rencontrés ou « ressentis » divers des utilisateurs/clients. Le chef de produit réalise cela dans l'urgence car le processus de conception complet est soumis à d'importantes contraintes de temps : ce renouvellement doit accompagner, voire anticiper les dernières technologies sur le marché (nouvelle génération de navigateurs, nouveaux fonctionnements de produits/services de la

concurrence) et, plus ponctuellement, observer des dates de lancement du produit. Une fois le cahier des charges réalisé, il est transmis à un ingénieur qui s'occupe de gérer le développement informatique nécessaire soit en interne dans la branche opérationnelle, soit en faisant appel à des concepteurs de FTA ou à d'autres entreprises. Les concepteurs sont choisis en dernier lieu dans ce processus et sont seulement les artisans de la réalisation spécifiée.

De plus, lors de l'entretien avec l'autre chef de produit, celui-ci exprime l'idée que trop d'études amont à la conception sont inefficaces : *« Il vaut mieux sortir quelque chose, on dit : bon, on fait le choix là, on arrive à un produit ou un service, on regarde et on rectifie ce qui ne va pas. Parce que vouloir avoir une approche trop construite dans laquelle on essaye de rationaliser, d'objectiver chacun des choix, je ne pense pas que ce soit compatible avec l'état du marché où il faut aller très vite d'abord. Et, en fait, en essayant de raisonner trop sur chaque choix, on perd un peu la cohérence du produit que peut avoir justement un responsable produit, en disant : et bien non, ma vision c'est ça. A la limite on relance le produit et puis, au moins, on le termine et on voit si c'est viable ou pas. C'est plus à la limite un processus de feedback plutôt que de mettre en œuvre une méthodologie sophistiquée. [...] Franchement, moi je pense que si je sors un produit, je demande des tests, qu'on me dise, simplement comment les gens réagissent. [...] On n'a pas toujours le temps de faire... Par exemple, on peut avoir beaucoup de chiffres, ce n'est pas toujours facile, mais déjà avoir des indicateurs, par exemple : quelles sont les pages qui seront regardées, quelles sont les pages qui ne sont pas regardées, pourquoi. Et tout ça c'est plus des outils a posteriori, plutôt que des outils a priori. »* A partir de cela, nous faisons aussi l'hypothèse que faire une étude approfondie des besoins en connaissances de chaque professionnel contribuant aux transformations ou les réalisant, semblerait encore plus vaine aux chefs de produit.

Malgré le précédent audit très orienté veille technologique et besoins clientèle ni les chefs de produit ni davantage les professionnels de l'évaluation n'ont l'idée ou la possibilité d'analyser en profondeur les besoins en connaissances des divers professionnels qui vont contribuer à transformer le produit/service et, plus particulièrement, ceux des concepteurs. Le demandeur de l'étude et l'utilisateur des résultats sont souvent des acteurs différents dans l'entreprise (Daniellou, 2004 ; Roussel, 1996). Il semble en être de même concernant les deux situations précédentes et les évaluations en phase de « feed-back » commandées au service d'évaluation. La question de l'utilité et de l'utilisation des résultats n'est donc pas observée.

3.3. LA QUESTION DE LA « VERITE » ET DE L'UTILITE D'UNE EVALUATION

Il s'agit, selon les professionnels de l'évaluation, d'être au plus proche des catégories de l'utilisateur pour créer le moins de distorsion d'avec l'activité perceptive en situation naturelle que l'on cherche à connaître. Nul doute que la formulation du questionnaire doit être la plus proche de l'entendement du répondant. Il s'agit bien d'éviter les décalages de compréhension entre le concepteur de l'outil d'évaluation et le sujet qui répond.

Cela concerne la formulation des questions à poser aux répondants, or il est possible de dissocier, en partie au moins, ce sur quoi doit porter l'évaluation et la formulation des questions elles-mêmes. Concernant ce sur quoi doit porter l'évaluation, « coller » le plus possible aux catégories des utilisateurs/clients amène aussi à passer à côté de la question de l'utilité des résultats produits. A quoi, à qui doivent-ils servir ? A partir de ces résultats, quelles décisions peuvent être réellement prises ? Sur quelles caractéristiques ou artefacts des produits/services, les professionnels peuvent-ils véritablement agir afin de les corriger et de mieux prendre en compte l'utilisateur ou le consommateur ? Quelle pertinence présente, par exemple, un résultat d'évaluation qui ne peut donner lieu à une transformation ? Le résultat en question demeurera une connaissance certes, mais inutilisée. Par exemple, apprendre que les femmes sont celles qui consultent le plus la présentation multimédia détaillant les fonctionnalités d'une application logicielle pour l'accès à Internet avant son installation, et les hommes le moins, constitue une connaissance qui n'apporte toutefois pas d'éléments assez précis aux concepteurs pour entreprendre une action d'amélioration¹²⁴.

En ce sens, les professionnels de l'évaluation produisent des connaissances qui ne prennent pas en compte la réalité des professionnels qui modifient le produit/service, de leurs représentations et des exigences de leur activité (arbitrage de contraintes, marges de manœuvre réelles). Il s'agit d'un mouvement épistémique sans prendre en compte la question de l'utilité des résultats. Ce point de vue est en étroit lien avec le contexte général dans lequel l'activité d'évaluation s'inscrit. On ne saurait dire cependant que l'orientation qui guide l'action des professionnels de l'évaluation est seulement une conséquence de ce contexte.

Dans ce qui suit, nous allons recueillir les réactions des professionnels de la transformation quant à la liste de dimensions utilisateurs et explorer leur vocabulaire et leurs représentations, puis les exigences de leur activité et leurs réels besoins relatifs à cette activité.

4. LA COMPREHENSION DES PROFESSIONNELS DE LA TRANSFORMATION QUANT A LA LISTE DE DIMENSIONS UTILISATEURS

4.1. CONDUITE ITERATIVE D'ENTRETIENS

Nous avons présenté la liste de dimensions utilisateurs concernant les sites Internet marchands (élaborée au chapitre précédent) aux quelques professionnels de la veille-marketing et de la conception et avons recueilli leurs réactions. Avec les deux professionnels de la veille-marketing, nous avons mené divers entretiens et exercices, décidés au fur et à mesure des données recueillies, puis un entretien avec chacun des deux ingénieurs multimédia de FTA et enfin un entretien de validation avec le responsable du Webmastering du Sitemarchand.com (rencontré par l'intermédiaire du professionnel de la veille-marketing spécialiste du Sitemarchand.com). L'entretien de validation a porté sur la liste de dimensions utilisateurs transformée à partir de certaines données des entretiens avec les quatre autres professionnels.

4.1.1. Avec les deux professionnels de la veille-marketing

- Des premiers entretiens libres à l'exercice de hiérarchisation :

Nous procédons d'abord à des entretiens libres (avec prise de notes et enregistrement mais retranscrits seulement dans un second temps), d'une à deux heures, sur le lieu de travail des professionnels de la veille-marketing, au cours desquels nous avons pu cerner les principaux objets de leur activité et leurs outils (outils de mesure d'audience, etc.). Le premier professionnel de la veille-marketing spécialiste du Sitemarchand.com nous énonce aussi spontanément une liste qui équivaut à ses besoins en évaluation. Avec le second professionnel de la veille-marketing, l'entretien est orienté directement sur la liste exhaustive des dimensions utilisateurs des sites Internet marchands (31 dimensions), car nous le connaissons déjà depuis l'étude sur les sites Internet portail. Le professionnel propose spontanément de former des regroupements de ces dimensions, très nombreuses selon lui, après quoi nous lui présentons la liste de sous-dimensions (plus d'une centaine). Pressé par le temps et inquiet de la difficulté de les appréhender toutes, il a exprimé l'envie de regarder cette liste ultérieurement, seul et à tête reposée.

N'ayant obtenu que des éléments d'ordre assez général durant ces premiers entretiens, nous espérons recueillir des informations plus précises en plaçant ces deux professionnels

seuls face à l'ensemble de ces dimensions et sous-dimensions, dans la situation hypothétique de recevoir des résultats d'évaluation selon chacune d'elles. L'objectif recherché est qu'ils indiquent, le plus systématiquement possible, l'importance de chacun de ces résultats dans leur activité. Nous optons donc pour un exercice de hiérarchisation.

- De l'exercice de hiérarchisation aux entretiens approfondis avec le Sitemarchand.com :

Nous leur envoyons une sorte de guide de test par messagerie électronique qu'ils nous retournent après avoir complété le document. Ce guide comporte la liste exhaustive des dimensions utilisateurs des sites Internet marchands ainsi que leurs sous-dimensions¹²⁵ avec une colonne vide pour que ces professionnels « notent » l'importance de ces diverses dimensions (ou plutôt résultats d'évaluation selon ces dimensions) dans leur activité de mise à l'épreuve des sites marchands. La notation se fait selon une échelle de Likert en 5 points et à items verbaux, à savoir « inutile », « peu important », « relativement important », « important », « très important ». Afin d'obtenir la hiérarchisation la plus fine possible, nous voulons utiliser une échelle à items verbaux puisque les sujets ont une préférence et facilité à utiliser ce type d'échelles plutôt que des échelles numériques (Urdapilleta, 2001c). Ils sont également invités à modifier les intitulés des dimensions et sous-dimensions si ceux-ci ne sont pas appropriés et à ajouter tout commentaire libre qu'ils jugent utile.

Etant donné le caractère en réalité abstrait et fastidieux de cet exercice, nous ne recueillons aucun commentaire ou modification des intitulés des dimensions. Pourtant, nous apprenons, lors d'échanges informels ultérieurs, qu'ils ne savent pas discerner avec certitude ce que certaines (sous-)dimensions contiennent, et d'autres leur apparaissent redondantes. Ils nous expliquent également que les sous-dimensions sont beaucoup trop nombreuses (plus d'une centaine) et qu'ils ne peuvent pas envisager l'utilité d'une telle quantité d'éléments ni, de ce fait, véritablement réaliser l'exercice de hiérarchisation. Ils s'y perdent.

Cela nous conduit à entreprendre un dernier travail d'élaboration avec ces professionnels. Malgré leurs vives critiques, une seconde chance nous est accordée et prend la forme d'un entretien semi directif, utilisant le Sitemarchand.com et la liste de dimensions sans leurs sous-dimensions (entretiens enregistrés et partiellement retranscrits à ce stade). Cette fois-ci, nous entamons la séance par une phase d'exploration rapide du site (20 minutes environ) qui fait office de mise en contexte et de support d'analyse. L'objectif ici n'est pas d'évaluer l'efficacité de la liste de dimensions dans ses éventuelles performances d'inspection d'utilisabilité tel que l'ont fait, par exemple, Bastien et Scapin ou Nielsen et Molich. Ce

support doit permettre aux professionnels, à tout moment de la discussion sur les dimensions, d'illustrer leurs propos d'exemples concrets. La séquence d'utilisation est aussi destinée à lever d'éventuelles incompréhensions relatives à des écarts de vocabulaire entre le chercheur et ces professionnels. Enfin, elle doit permettre de repérer à quels artefacts des sites marchands les professionnels attachent plus d'importance (avec prudence, car les professionnels deviennent dans cette situation eux-mêmes utilisateurs, voire clients potentiels). Pour le second professionnel, toujours plus pressé, cette séquence d'exploration, plus courte, a lieu à la fin de l'entretien par crainte de manquer de temps pour aborder toutes les dimensions.

Lors de ces entretiens, nous reprenons aussi l'intitulé de chacune des dimensions et leur demandons ce qu'ils en comprennent, s'il faut la formuler autrement et comment. Simultanément, nous leur demandons de se replacer dans la situation d'être en possession de différents résultats d'évaluation subjective et nous leur demandons si et en quoi ils leur seraient utiles dans leur activité de mise à l'épreuve des sites et de prise de décisions diverses concernant ceux-ci.

Rapidement, dans la discussion à caractère très libre, se fait à nouveau sentir la nécessité de réaliser des regroupements à partir de ces 31 dimensions. A nouveau, car ils ont déjà été suggérés et, en partie, élaborés par le second professionnel de la veille-marketing, lors de son premier entretien. Cette fois-ci, c'est le professionnel spécialiste du Sitemarchand.com qui les demande, et des ensembles sont formés verbalement et nommés par lui au fur et à mesure de son entretien. Ce même travail est repris avec le second professionnel de la veille-marketing, mais de façon plus systématique : nous lui proposons de construire et de nommer des ensembles à partir des numéros des dimensions sur une feuille libre. L'activité de catégorisation de ce second professionnel s'apparente à la technique de tri libre simple avec nomination des groupes formés (Bisseret *et al.*, 1999). Ils énoncent aussi, tous deux, en partie au moins, les raisons pour lesquelles ils réalisent les regroupements.

L'ordre d'enchaînement des « exercices » au sein de chaque entretien n'a pas d'importance, étant donné qu'il s'agit d'utiliser les données dans un processus de conception (d'un artefact pour ces professionnels).

4.1.2. Avec les deux ingénieurs : entretiens approfondis avec le Sitemarchand.com

Les mêmes entretiens (enregistrés et partiellement retranscrits) avec la phase d'exploration du Sitemarchand.com sont réalisés avec les deux ingénieurs multimédia. Ils s'adonnent, de même que les précédents professionnels, à divers commentaires critiques du site marchand, en tant qu'utilisateur qu'ils sont alors et aussi en tant qu'expert analysant l'interface et la prise en compte de l'utilisateur dans la conception du site Internet.

A la suite de cette première phase, plus ou moins longue pour l'un et l'autre, leur compréhension des intitulés de chacune des dimensions est vérifiée. Les sources de confusion et leurs propositions d'amélioration sont aussi précisées. Le travail de catégorisation n'a pas lieu car ils n'en expriment pas le besoin. En revanche, chaque sous-dimension des dimensions relatives à la recherche de produits vendus sur le site et au processus de commande (panier, formulaires de transaction, etc.) sont passées en revue. Notre objectif est d'obtenir davantage de données sur certains éléments relatifs à l'utilisation de l'interface, que les professionnels de la veille-marketing considèrent comme des acquis (minimums requis qui doivent être assurés par les concepteurs). Les professionnels de la veille-marketing sont, en effet, moins « loquaces » à propos de certains de ces éléments (ex. la recherche interne de produits sur le site) qui ne sont pas directement relatifs à leur champ d'action.

4.1.3. Avec le responsable du Webmastering du Sitemarchand.com : entretien de validation

Suite aux divers entretiens avec les précédents professionnels, nous nous apercevons des insuffisances de la liste de dimensions utilisateurs tant en termes de compréhension qu'au niveau de sa forme (liste cumulant 31 dimensions et ses sous-dimensions, partagée en seulement deux grands ensembles). Nous avons donc procédé à la transformation de surface de cette liste, explicitée dans ce qui suit. C'est donc muni de cette nouvelle liste que nous avons mené un entretien avec le responsable du Webmastering du Sitemarchand.com¹²⁶.

Nous avons fait valider chacune des dimensions et leur nouvelle définition ainsi que les regroupements qui les structurent et nous avons recueilli encore quelques rectifications.

Les divers entretiens sont récapitulés dans le tableau suivant :

Tableau 12 : Récapitulatif des entretiens menés avec les professionnels de la veille-marketing et de la conception, avant et après transformation de la liste exhaustive des dimensions utilisateurs

	1 ^{er} prof. ^{el} VM spécialiste MARCH	2 ^{ème} prof. ^{el} VM non spécialiste MARCH	2 ingénieurs FTA	resp. Webm. MARCH
Type de sujet	1 et 2	2	2	2
Premiers entretiens libres	✓ + liste de besoins en évaluation, énoncée spontanément	✓ + regroupements spontanés des dimensions		
Exercice de hiérarchisation des dimensions et sous-dimensions	✓	✓		
Entretiens avec utilisation du Sitemarchand	✓ + regroupement des dimensions	✓ + regroupement des dimensions	✓	
Transformation de surface de la liste exhaustive de dimensions	Réalisée à partir des précédents entretiens avec les deux professionnels de la veille-marketing et avec les deux concepteurs multimédia.			
Entretien de validation de la transformation				✓

Légende : VM : veille-marketing ; MARCH : entreprise Sitemarchand.com ; Webm. : service Webmastering ; ch. graph. : charte graphique ; mktg. fidél. : marketing de fidélisation ; sujets type 1 : ont une activité d'évaluation ; sujets type 2 : ont une activité de conception/marketing des produits/systèmes

4.2. LA TRANSFORMATION DE SURFACE DE LA LISTE EXHAUSTIVE DE DIMENSIONS UTILISATEURS

Nous rendons compte ici de la transformation opérée à partir d'une première analyse sommaire des précédents entretiens avec les professionnels de la veille-marketing et les ingénieurs. Dans un premier temps, les remarques des quatre professionnels nous ont fait comprendre que la liste exhaustive de dimensions utilisateurs devait évoluer à deux niveaux, dans le sens :

- 1/ de favoriser une lecture et une compréhension plus efficace,
- 2/ d'aider la compréhension au niveau du choix du vocabulaire.

Concernant le premier point, nous avons procédé à :

- l'élaboration d'une définition pour chaque dimension synthétisant toutes leurs sous-dimensions (le format de l'artefact de départ est donc modifié),
- et au regroupement des dimensions dans des grands ensembles permettant de repérer des thèmes principaux avant de se lancer dans la découverte des 31 dimensions.

Concernant le second point, nous avons procédé à :

- la modification des intitulés des dimensions pour éviter toute source de confusion,

- et l'ajout d'éléments de leur langage professionnel pour qu'ils reconnaissent plus rapidement ce dont il s'agit.

Voici le tableau récapitulatif des entretiens utilisés pour les différentes opérations de cette transformation :

Tableau 13 : Récapitulatif des entretiens utilisés pour les différentes opérations de la transformation de surface de la liste de dimensions utilisateurs des sites Internet marchands

	Premiers entretiens libres, 2 professionnels VM	Entretiens avec le Sitemarchand.com, 2 professionnels VM	Entretiens avec le Sitemarchand.com, 2 ccpteurs multimédia
Elaboration d'une définition pour chaque dimension (nouveau format d'artefact)	(✓)	✓	✓
Regroupement des dimensions	✓	✓	
Modification des intitulés des dimensions		✓	✓
Ajout d'éléments de leur langage professionnel		✓	

Nous abordons ci-après le détail de la transformation de l'artefact. La liste de dimensions utilisateurs après sa transformation de surface est présentée en intégralité en annexes du chapitre 4¹²⁷.

4.2.1. Un nouveau concept d'artefact pour favoriser une lecture et une compréhension plus efficaces

4.2.1.1. Un concept d'artefact au format intermédiaire

Suite à l'exercice de hiérarchisation où les professionnels de la veille-marketing ont vivement critiqué le nombre de sous-dimensions, lors des entretiens avec le site Sitemarchand.com, nous leur présentons la liste avec les dimensions seulement. Cependant, pour la plupart d'entre elles et concernant les quatre professionnels, nous sommes amené à faire un court résumé de ce qu'elles contiennent. Celui-ci est parfois inutile car évident et parfois il permet de lever des discordances déjà amorcées, liées à une interprétation de leur part en décalage avec ce que la dimension désigne réellement.

En fait, l'artefact avec les seuls intitulés des dimensions est insuffisant et celui avec toutes les sous-dimensions trop laborieux. Un format intermédiaire s'impose. Un nouveau

concept a été élaboré en reformulant l'ensemble des sous-dimensions dans chacune des dimensions sous la forme d'une définition plus synthétique mais néanmoins complète¹²⁸.

4.2.1.2. Le regroupement des dimensions pour structurer cet artefact

Seuls les professionnels de la veille-marketing, mais lors des deux types d'entretiens (libres, puis avec le Sitemarchand.com) « réclament » spontanément des grands ensembles regroupant ces 31 dimensions trop nombreuses à leurs yeux. Pour élaborer ces grands ensembles, nous nous sommes basé sur les données suivantes :

- Liste sommaire énoncée spontanément par le premier professionnel de la veille-marketing, lors du premier entretien :
 - Produits (choix de) : répondre à la manière dont les clients réagissent,
 - Conditions de mise en vente (livraison, retours de produits, facilités de paiement),
 - Prix,
 - Aide à la sélection de produits,
 - Explications / Informations sur les produits (ex. Newsletter).

- Regroupements verbalisés spontanément par le second professionnel de la veille-marketing à partir des dimensions présentées, lors du premier entretien :
 - Confiance / Notoriété (*« Ca va donner le cadre, vous entrez chez Darty, vous entrez chez France Télécom, vous entrez chez l'inconnu du coin...On a déjà un cadre implicite de sécurité, de confiance, de savoir où on va quoi. C'est comme si tu rentres sur l'autoroute ou tu pars sur un petit chemin de campagne »*),
 - Offre (choix de produits proposés),
 - Détails sur les prix, sur l'offre, sur les produits (informations données),
 - Sécurité, mode de paiement,
 - Utilisabilité, présentation, page d'accueil, navigation (par analogie au magasin : *« c'est le rayonnage »*, la facilité à trouver le produit dans un rayon),
 - Relation vendeur-client, S.A.V.,
 - Moteur de recherche, conseils, aides à la vente (comme la *« signalétique »* dans un grand magasin - mercerie, droguerie, quincaillerie, etc. - et le *« petit prospectus »* au bout de chaque linéaire ou rayon pour annoncer les produits les moins chers, etc.),
 - Rapidité de chargement.

- Regroupements réalisés verbalement par le premier professionnel de la veille-marketing à partir des dimensions présentées, lors de l'entretien avec le Sitemarchand.com :

- Notoriété,
- Facilité d'accès,
- Accueil,
- Offre,
- Communication et visibilité des « choses » & Informations produits,
- Facilité à trouver les produits,
- Panier & Commande,
- Sécurité,
- S.A.V.
- Incontournables (les huit dernières dimensions dans la liste étaient déjà regroupées sous l'intitulé « dimensions génériques à tout site Internet »).

- Regroupements réalisés par écrit par le second professionnel de la veille-marketing à partir des dimensions présentées, lors de l'entretien avec le Sitemarchand.com :

- *"Le client arrive sur le site"*,
- *"Le client est sur le site"*,
- *"Le client commande"*,
- *"Le client achète"*,
- *"Le client reçoit le produit"*,
- et les huit dernières dimensions dans la liste sont laissées de côté (les « dimensions génériques à tout site Internet »).

La logique des dimensions utilisateurs présentées lors des entretiens suit le cheminement de l'internaute réalisant une commande par Internet jusqu'à réception du produit. Dans l'ensemble, cette logique est peu modifiée par les professionnels, même s'ils déplacent quelques dimensions au sein des grands ensembles qu'ils forment. Ainsi, ces quatre listes issues des différents entretiens présentent de nombreux éléments similaires ou qui se recourent¹²⁹. En établissant les correspondances entre les éléments des quatre listes tout en se basant davantage sur les deux dernières élaborées de façon plus systématique, nous avons pu dégager neuf thèmes :

- « Avant l'arrivée sur le site » (notoriété du site, facilité d'accès),
- « Accueil » (page d'accueil, contact vendeur, etc.),
- « Offre ou positionnement du site » (choix de produits, prix, etc.),

- « Communication et valorisation de l'offre » (détails des produits, envois d'informations personnalisées comme des « Newsletters », etc.),
- « Recherche » (champ de recherche par mots clefs, etc.),
- « Commande » (simplicité du panier, etc.),
- « Achat » (sécurité, modes de paiements, etc.),
- « Livraison et S.A.V. » (rapidité de la livraison, contact S.A.V., etc.)
- et enfin les « Incontournables » (rapidité de chargement, esthétique, présentation des informations, mise à jour, etc.).

Ces neuf thèmes ne sont pas définitifs. Le but est de proposer des ensembles permettant déjà de rompre avec cette impression d'accumulation trop exhaustive et de permettre d'appréhender plus facilement les 31 dimensions : les professionnels doivent savoir à quel grand thème chaque dimension se rapporte, avant de la lire.

4.2.2. Modification de l'artefact au niveau du vocabulaire

4.2.2.1. Les sources de confusion entre dimensions et modification des intitulés

Les modifications réalisées en terme de vocabulaire à partir des verbalisations des professionnels de la veille-marketing (recueillies lors de l'entretien avec le Sitemarchand.com) sont les suivantes :

- les « informations de vente » qui peuvent être confondues avec les « informations sur les produits », elles ont été remplacées par les « conditions générales de vente »,
- l'imprécision constituée par les « fonctions de recherche » a été transformée en « différentes possibilités de recherche de produits ».

Les concepteurs, quant à eux, nous suggèrent des dénominations plus explicites comme :

- « bons de commande » au lieu de « la commande » qui est moins spécifique,
- et « preuve » et « confirmation » de « l'opération de commande en ligne » en remplacement du trop ergonomique « feed-back sur la commande »,
- ou encore « sécurisation et différents modes de paiement » plutôt que le plus laconique « sécurité et mode de paiement ».

Enfin, d'autres dimensions par leur formulation trop imprécise sont source d'ambiguïté pour la plupart des quatre professionnels et elles ont été transformées. Les raisons de ces modifications sont synthétisées dans le tableau ci-dessous :

Tableau 14 : Récapitulatif des intitulés des dimensions des sites Internet marchands sources de confusion et explications des modifications nécessaires

Intitulé initial des dimensions :	Ce qui est source de confusion :
« Facilité d'accès (adresse, référencement) »	Il est nécessaire de préciser qu'il s'agit de la facilité d'accès au site sur le « plan technique » par un outil de recherche sur le Web ou en saisissant une adresse URL. Ainsi, la confusion ne sera plus possible avec le référencement « commercial » par des sites Internet de comparaison des prix.
« Clarté du service rendu et informations qui rassurent dès la page d'accueil »	Il faut clarifier « le service rendu » qui désigne tout ce que le site offre à la vente (produits et services), en plus des services de livraison ou d'après vente. « Le service rendu » est en effet lu par certains comme désignant uniquement les services après vente.
« Existence et qualité d'une relation humaine au-delà de la machine (conseiller en vente) »	Il faut préciser que cette dimension désigne seulement le fait de pouvoir contacter un vendeur pour une aide ou des conseils pour l'achat. Le fait de pouvoir contacter un livreur ou le service clients pour des problèmes de paiement, de produits défectueux, etc. concerne une autre dimension.
« Différences d'Internet par rapport au commerce classique (même choix de prix, produits, promotions) »	Il faut préciser : différence entre un site Internet et le magasin de la même enseigne. De plus, le premier professionnel de la veille-marketing nous a fait remarquer que les clients font des comparaisons entre tout site Internet et tous magasins de proximité du même secteur de vente. La dimension a donc été généralisée.
« Informations simultanées sur la sécurité/vente et feed-back sur la commande »	Une confusion a lieu entre le « feed-back sur la commande » (confirmation de la commande sur le site ou par courriel) et le « suivi de la commande » (courriels suivant la confirmation ou messages sur le panneau de suivi de commande sur le site). Cet élément a été reformulé.
« Possibilité de pré-commande, suivi de la commande »	La « pré-commande » est entendue comme le simple fait de pouvoir commander un produit proposé en avant-première. Or cela désigne la fonction pour retenir un produit inexistant en stock et être prévenu de sa prochaine remise en vente (système d'alertes). Les deux éléments ont été précisés.

4.2.2.2. Ajout d'éléments du langage professionnel

Des mots issus du langage technique du « marketing en ligne » relevés dans les verbalisations des deux professionnels de la veille-marketing ont enfin été ajoutés aux intitulés et définitions des dimensions. L'objectif est que ces professionnels les identifient plus facilement grâce à un vocabulaire plus familier. Concernant les ingénieurs multimédia, cela n'a pas été nécessaire.

Nous avons ajouté des termes comme :

- positionnement payant ou « pôle position » indiquant que le référencement du site sur les outils de recherche sur le Web peut être déterminé moyennant financement,
- « Cross-Selling » qui désigne le fait de proposer d'autres produits similaires ou accessoires complémentaires à l'un de ceux qui a été sélectionné par un internaute (ex. « les internautes ayant acheté ce produit ont aussi aimé... »),

- les termes « Opt-in » et « Opt-out » qui tiennent au fait qu'une option soit posée ou non concernant les informations susceptibles d'être envoyées à l'internaute (ex. la case « je ne désire pas recevoir d'informations complémentaires des partenaires de l'entreprise » cochée par défaut ou non). Selon le cas, l'internaute doit cocher ou non la case pour ne pas recevoir de courriels (ou l'inverse).

Après avoir donné à voir la transformation de surface opérée sur la liste de dimensions, nous évoquons son insuffisance.

4.3. INSUFFISANCE DE LA TRANSFORMATION DE SURFACE

Dans un premier temps, nous pensions que la précédente transformation utilisant leur vocabulaire et, en partie, leurs représentations (regroupements des dimensions) était utile pour eux. Cependant au fur et à mesure de la réalisation de cette dernière, certaines remarques, du professionnel de la veille-marketing spécialiste du Sitemarchand.com notamment, nous font douter, non pas de la plus grande facilité à appréhender le nouvel artefact mais de l'utilité même de cet artefact, transformé ou non, dans leur activité.

Nous avons alors retranscrit systématiquement l'ensemble des entretiens enregistrés des deux professionnels de la veille-marketing qui sont plus critiques et plus concernés par la question de l'évaluation (seul le premier entretien libre avec le premier professionnel de la veille-marketing, pour des raisons techniques, n'a pas pu être enregistré). Nous avons ainsi identifié les données, c'est-à-dire les verbalisations, nous permettant de penser que cet artefact transformé ou non ne les aide pas à penser, à concevoir l'évaluation d'un site Internet marchand.

4.3.1. Fournir concrètement ce sur quoi l'évaluation d'un site marchand peut porter

Premièrement, le professionnel de la veille-marketing, spécialiste du Sitemarchand.com, lors de l'entretien avec ce site, évoque le fait que l'intitulé des dimensions (non transformées) ne permet pas de connaître ce qu'il s'agit précisément d'évaluer. Concernant la dimension « Existence et qualité de la relation humaine au-delà de la machine » (devenue ensuite « Existence et qualité d'un contact avec un vendeur ») : « Si tu poses la question à quelqu'un qui analyse un site finalement, enfin un client quoi. Donc c'est existence. Est-ce qu'elle existe ? Est-ce que vous la voyez ? D'abord est-ce que vous la souhaitez ? Est-ce qu'il la

souhaite finalement ? Est-ce qu'il en a besoin en général ou pas ? Après, il peut ne pas en avoir besoin même si elle existe, même si elle est disponible. Et après c'est donc : [...] Est-ce que vous l'avez perçue, c'est-à-dire est-ce qu'elle existe sur le site que vous voyez ? Il peut très bien dire non alors qu'elle existe, c'est-à-dire qu'il ne la voit pas. Et après il peut très bien l'avoir vue et puis elle peut ne pas être de bonne qualité, c'est-à-dire le service-client ou la hotline peut être défaillante. » Nous voyons là que « Existence » et « qualité » dans l'intitulé de la dimension sont censés refléter les éléments détaillés par le professionnel qui pourtant trouve cet intitulé insuffisant puisque justement il éprouve le besoin de l'explicitier en détail. Il fait cela à plusieurs reprises concernant d'autres dimensions.

C'est cela qui nous avait conduit, dans la transformation précédente, à rédiger des définitions synthétisant toutes leurs sous-dimensions (censées indiquer plus en détail les éléments possibles d'une évaluation). Cependant, l'exemple précédant et ce qui suit plus loin nous font comprendre que ces professionnels n'ont pas besoin qu'on leur définisse des dimensions, nécessairement trop généralisatrices, mais qu'on leur fournisse concrètement ce sur quoi l'évaluation d'un site marchand peut porter.

4.3.2. Préciser les objectifs de l'évaluation liés aux caractéristiques du produit/service

De plus, se poser concrètement la question de ce que l'on a besoin d'évaluer concernant un site marchand donné nécessite de le replacer dans son contexte (positionnement par rapport à la concurrence, phases d'évolution du site, etc.). Cela déterminera l'objectif de l'évaluation. Le professionnel de la veille-marketing, spécialiste du Sitemarchand.com, nous l'explique : « *En fait, ce qui se passe c'est qu'un site comme [...], par exemple, puisqu'on a celui-là sous les yeux, en fait, il peut avoir un positionnement et un objectif. Il peut se dire je vais vendre le maximum de choses possibles sur les livres, sur les CD. Je veux être exhaustif sur l'ensemble des produits. Il peut y en avoir d'autres qui vont viser soit une clientèle différente, un segment clientèle, soit faire une offre limitée à un sous-ensemble de ce qui est possible [...]. Voilà parce qu'ils ont une autre cible et parce qu'ils ont un positionnement différent. La question est de savoir qu'est-ce qu'on veut évaluer. Est-ce qu'on veut évaluer... Est-ce que le client perçoit bien ce qu'on a voulu faire ? Ou est-ce qu'on veut [faire] évoluer notre site en fonction de l'évolution de notre clientèle ? [...] On peut peut-être quand même se fixer deux contextes. Contexte de ce que l'on veut, ce qu'on a voulu faire est-il bien perçu par le client ? Et deuxième axe : on veut changer, donc on veut savoir notre cible clientèle ce qu'elle souhaite. Restons sur le premier cas, si c'est pas bien perçu c'est qu'on n'a pas créé*

le site correctement, qu'on n'a pas fait les bonnes..., que la communication, la visibilité des choses n'est pas bonne. On n'a pas envie pour autant de changer notre positionnement. Dans le deuxième cas : on est en train de chercher... On a défini notre cible et on veut s'adapter à notre cible. [...] Mais dans ce cas là, effectivement, ce qu'on veut percevoir c'est quelles sont les attentes ? ».

En reprenant le même thème, ce professionnel exprime de façon plus évidente encore que les dimensions seules sont toujours insuffisantes et qu'elles doivent être replacées dans le contexte d'évaluation qui doit être précisé. Ce dernier est lié notamment au processus de développement du site : *« Je ne comprends pas qu'on puisse poser cette phrase là (intitulé de la dimension), indépendamment de ce qu'on a dit tout à l'heure, de la définition d'un axe d'enquête. Il y a deux axes d'enquête qu'on a retenu tout à l'heure pour la même question, pour une question préalable du moins. [...] D'ailleurs je devrais le dire dans l'autre sens : la cible clients que l'on vise, quelles sont ses attentes, on fait un site qui correspond aux attentes, ensuite on regarde si ce qu'on a fait est bien perçu. Donc il y a un processus d'approche de la vente et du client qui me semble trop concaténé là ici, et pas suffisamment éclaté pour qu'on perçoive bien les différentes nuances qu'il peut y avoir. »*

Le second professionnel de la veille-marketing a aussi soulevé la difficulté de travailler à partir de ces dimensions qui sont hors contexte et la nécessité de les replacer dans le processus de conception avec les différentes phases d'évolution du site Internet : *« Si tu veux que je fasse des hiérarchies à ce point, parce que ce n'est pas évident. [...] Ce n'est pas évident de choisir non plus, parce que... Et puis ça dépend à quelle phase on en est. [...] Oui et puis où tu en es de ta conception même de ton site Web là. Si tu es en reformulation, si tu es en refonte de site, si tu es en... »*. Nous apprenons dans la suite de l'entretien que les phases de refonte ou de reformulation ne sont pas spécifiquement définies dans son esprit, mais l'idée est bien là aussi.

4.3.3. Opérationnaliser les dimensions pour constituer un outil de conception d'une évaluation

Les propos des professionnels de la veille-marketing nous avaient mis sur la voie de la création de grands ensembles pour qu'ils appréhendent plus rapidement les dimensions : *« Et bon c'est pour ça, que j'avais besoin de regrouper, à quoi ça a attiré, ça a attiré à l'accueil, ça a attiré à l'offre. Et en même temps, ici, donc c'est pas forcément regrouper mais*

éclater... ». Des déplacements de quelques dimensions au sein de ces ensembles avaient aussi été assurés (pour répondre à « *J'ai envie soit d'éclater, soit de concaténer différemment* »).

Néanmoins, le problème de fond est d'une autre nature : hormis l'absence de précision du contexte, les dimensions sont perçues comme des recommandations à suivre et ne constituent pas des éléments permettant de construire une évaluation. « *...la difficulté que j'ai, c'est effectivement d'être toujours en train de me dire : je ne vois pas ce qu'on a derrière, je ne peux pas regrouper les choses comme ça. [...] C'est pour ça qu'ici je voulais faire apparaître l'accueil, l'offre ensuite la recherche effectivement. Et j'ai l'impression déjà que si tu guides les gens que tu veux conseiller finalement, parce qu'au bout du compte c'est ça que tu veux faire. Tu veux faire une thèse pourquoi ? Pour conseiller des marketers et des concepteurs de site. Tu veux faire un manuel. C'est quoi ton objectif ? [quelques explications trop succinctes...]* Donc c'est bien pour nous que tu fais un manuel ou enfin quelque part une liste de recommandations ou un mode d'emploi. » L'objectif dans notre étude n'est pas de produire un manuel de recommandations.

L'objectif à ce stade n'est pas atteint. L'outil qui permettrait de les aider dans leur activité d'élaboration d'évaluations de sites marchands est encore à concevoir. La transformation des intitulés, la rédaction de définitions, le regroupement des dimensions en grands ensembles les aidera certes à mieux appréhender la liste de dimensions, mais l'artefact même transformé ne constitue pas un outil d'aide à la conception d'évaluations.

De même, lors de l'entretien de validation, le responsable du Webmastering nous exprime que la plupart des dimensions examinées à partir de la liste transformée correspondent à ses propres connaissances des internautes et consommateurs de la cible clients du Sitemarchand.com (obtenues à partir des remontées clientèles et d'autres études). Il reconnaît lui aussi cet artefact transformé comme une liste de recommandations, même s'il n'est pas dans la même problématique (son activité est davantage de concevoir que d'évaluer). Il faut préciser que les dimensions utilisateurs des sites Internet marchands sont exprimées de façon à retranscrire les jugements, les exigences, les préférences, etc. des utilisateurs/clients.

Ils nous font donc rompre avec la logique d'une liste de dimensions utilisateurs, artefact conçu par les professionnels de l'évaluation, et réfléchir à la conception d'un véritable outil, c'est-à-dire d'un artefact instrumentalisable par les professionnels du marketing (et de la conception) pour leur propre activité de construction d'évaluations. Il faut opérationnaliser ces dimensions trop analytiques.

5. DEUX MOUVEMENTS EPISTEMOLOGIQUES DISTINCTS

Si les professionnels de l'évaluation n'ont pas de mal à transposer les dimensions utilisateurs sous la forme de dimensions à évaluer pour élaborer une évaluation (processus de médiations pragmatiques), les professionnels de la veille-marketing émettent de réelles réticences à réaliser un tel exercice.

La question à ce stade est la suivante : est-ce qu'il s'agit de conclure simplement que les professionnels de la veille-marketing ont besoin d'être « accompagnés » dans un processus partant des dimensions utilisateurs pour aboutir à la conception d'une évaluation alors que les professionnels de l'évaluation ont déjà conçu ce processus de transformation ou bien les deux types de professionnels sont-ils engagés dans deux mouvements épistémologiques bien distincts ?

Nous constatons, tout d'abord, que plusieurs éléments de l'activité des professionnels de la veille-marketing et aussi des professionnels du marketing du Sitemarchand.com sont bien sûr distincts de l'activité des professionnels de l'évaluation.

5.1. DES SITUATIONS, FINALITES ET COMPETENCES DIFFERENTES

Comme nous l'avons déjà vu, les professionnels de l'évaluation élaborent les dimensions utilisateurs en lien avec leur point de vue (approcher au plus près l'activité perceptive réelle des utilisateurs) et/ou en lien avec les objectifs donnés par des demandes d'évaluations peu spécifiées ; elles-mêmes certainement liées au cloisonnement dans l'organisation fonctionnelle de France Télécom. Au contraire, les professionnels de la veille-marketing et aussi du marketing, voire du Webmastering du Sitemarchand.com (comme nous le verrons aux chapitres suivants) sont engagés dans une activité d'évaluation d'un produit/service précis. L'évaluation porte alors sur un commerce électronique ayant des caractéristiques données au niveau de son interface, des capacités technico-commerciales plus ou moins arrêtées et des spécificités « marketing » correspondant à sa cible clients. L'évaluation dépend de plusieurs éléments que l'on peut spécifier avec précision à partir des précédentes caractéristiques données au moment de l'évaluation et des insatisfactions ou difficultés des utilisateurs/clients déjà pressenties (lors de précédents résultats d'évaluation, des réclamations au niveau du service clients, d'une expertise comparative avec des sites concurrents, etc.). Bref, les objectifs de l'étude sont très contextualisés au contraire de ceux de la situation des professionnels de l'évaluation. Leurs objectifs sont aussi plus finalisés. Ils évaluent pour

vérifier que leur produit apporte satisfaction (suite à une éventuelle transformation) et, le cas échéant, amorcer une action de correction. Les finalités des activités des uns et des autres ne sont pas les mêmes.

Enfin, les métiers ou compétences des deux types d'acteurs ne sont pas semblables non plus. Les professionnels de l'évaluation, issus d'une formation en ergonomie, en psycho-acoustique, en linguistique ou en statistique, sont centrés sur l'étude et l'élaboration de méthodes d'évaluation ou de traitement de données (ex. outils de « datamining » pour analyser des contenus déposés par des internautes sur des forums ou chats de sites Internet). Les professionnels formés au marketing ont une activité plus opérationnelle et sont centrés sur le recueil de données leur permettant directement d'agir pour augmenter les taux de vente sur leur site marchand, puisque telles sont leurs finalité et spécificité. Leurs connaissances en matière de méthodologies et de types de résultats qu'il est possible de produire sont certainement moins étendues, puisqu'ils n'ont recours qu'à quelques outils de recueil de données (panels¹³⁰, enquête par questionnaire, parfois *focus group*), probablement ceux dont ils ont le plus besoin. Enfin, les uns sont isolés de l'activité de transformation des produits/services au sens large (animation commerciale, conception, etc.), et pour les autres, il s'agit de leur activité principale.

5.2. DEUX MOUVEMENTS EPISTEMOLOGIQUES POURTANT DISTINCTS

Cela pourrait nous amener à répondre à la question précédente par le seul fait des finalités et compétences différentes des deux situations et types de professionnels : les professionnels de la veille-marketing ne parviennent pas à envisager, à concevoir une évaluation à partir des dimensions utilisateurs car ils ne sont pas impliqués dans la même situation, n'y ne possèdent les mêmes compétences.

Cependant la finalité des professionnels de l'évaluation, comme nous l'avons déjà exprimé, est aussi de produire les avis des utilisateurs/clients sur leurs « vrais » critères. Cela les situe dans un mouvement de production de connaissances, similaire à une démarche scientifique où l'agir est subordonné à la connaissance. Il s'agit alors de produire, d'enrichir, bref d'agir sur la connaissance.

Les professionnels du marketing et de la conception, quant à eux, sont engagés dans un mouvement opérationnel, pragmatique. Leur souci est celui de l'efficacité et de l'utilité : la production de connaissances est pour eux subordonnée à l'agir. Ils ont besoin que soient

produites les connaissances dont ils peuvent effectivement se servir. Nous avons vu qu'ils réagissent aux dimensions utilisateurs présentées en fonction de ce qu'ils savent avoir besoin d'évaluer pour produire les connaissances nécessaires pour agir. Nous faisons donc l'hypothèse que corrélativement aux situations (finalités, compétences) différentes, les postulats ou mouvements épistémologiques ne sont pas les mêmes.

Ainsi, nous proposons de partir de l'activité des professionnels transformant les produits et de chercher à connaître quels sont leurs réels besoins en connaissances pour piloter les actions commerciales et la (re)conception de leurs produits/services.

Dans cette perspective, il ne s'agit pas pour autant d'exclure une contrainte « côté utilisateurs » (voire « clients ») qu'ils ne voudraient pas ou auraient des difficultés à prendre en compte. Il s'agit d'apprendre à mieux connaître la réalité de leur activité aussi bien au niveau de leurs contraintes que de leurs connaissances et représentations pour produire des résultats adaptés, tout en ne sacrifiant pas des éléments relatifs aux attentes des cyber-consommateurs et internautes. S'arrêter à l'idée que ces professionnels n'en tiennent pas compte est une erreur. Il s'agit plutôt de réfléchir à une solution d'évaluation à format variable pour s'adapter aux exigences de leur activité et d'explicitier ces solutions convenablement.

Il s'agit de produire des artefacts non pas transformables ou interprétables en termes de contenus ergonomiques (ou autres contenus issus d'évaluations subjectives) mais instrumentalisables par eux. L'optique choisie ici est de refuser le principe selon lequel il s'agit simplement de produire des résultats utilisateurs amenés aux destinataires comme des vérités à prendre en compte, sans considération approfondie de la réalité de l'activité de ces derniers.

Par ailleurs, nous aborderons, aux chapitres 6 et 7, certaines limites ou difficultés posées par cette seconde optique, puisque des divergences existent entre le marketing et l'ergonomie. Nous évoquerons aussi le rôle de l'ergonome (ex. aider les professionnels à prendre conscience de certains réels « besoins en évaluation » qu'ils n'osent ou ne peuvent pas considérer).

6. CONCLUSION

Nous avons appris que le point de vue des professionnels de l'évaluation cherchant à révéler la « vérité » des utilisateurs est, à la fois, propre à ces acteurs et, à la fois, dépendant

des facteurs contextuels qui impriment leur activité. La prise en compte de facteurs économiques et organisationnels à l'échelle de l'entreprise permet, en effet, de resituer le processus de construction de l'évaluation et le point de vue de ces professionnels dans la dynamique plus large « demande d'évaluation – évaluation - transformation des produits/services » impliquant différents services et acteurs indépendants les uns des autres. Toujours est-il que les professionnels de l'évaluation n'ont aucun mal à transposer les dimensions extraites du matériau utilisateur en dimensions d'évaluation. Ils les instrumentalisent pour construire le questionnaire d'évaluation. Nous avons vu, en effet, les différentes médiations pragmatiques leur permettant de transformer ces dimensions en questionnaire, puis d'aboutir aux résultats restitués aux destinataires et demandeurs des études.

Dans cette situation, les professionnels de l'évaluation se construisent un instrument d'élaboration de leurs évaluations à partir de cet artefact « dimensions utilisateurs ». En revanche, dans la seconde situation, les professionnels de la veille-marketing (et par extrapolation, tout professionnel du marketing) ne parviennent pas à instrumentaliser cet artefact, alors qu'ils réalisent eux-mêmes des évaluations de leurs produits/services. Dans ce cas, l'artefact ne peut devenir un instrument. Les dimensions utilisateurs, même partiellement modifiées, restent inutilisables par eux.

Nous avons mené des entretiens qui visaient à produire des données pour réaliser une transformation de surface devant améliorer leur compréhension de ces dimensions utilisateurs. Ces entretiens nous révèlent progressivement que les dimensions doivent être opérationnalisées de façon à devenir un réel outil de construction des évaluations pour ces derniers. En effet, tel est leur besoin, lié à leur propre activité d'évaluation des produits/services qu'ils contribuent à concevoir et transformer, même s'ils s'occupent davantage de la dimension commerciale du projet. Leur approche de l'évaluation est plus utilitaire et plus contextuelle. Ils produisent les résultats, révélant les jugements des utilisateurs/clients, mais portant sur les caractéristiques de leur produit/service qu'ils savent devoir évaluer. Une connaissance de l'historique du commerce électronique (ex. évolution fonctionnelle de l'interface, investissement dans une machine de colisage supplémentaire, etc.) et de l'évolution de la clientèle et de sa perception leur permet cela.

A travers nos entretiens, nous découvrons qu'ils n'ont donc pas besoin de dimensions, formalisées et génériques, mais de voir concrètement ce sur quoi peut porter l'évaluation du

produit/service ainsi que l'objectif de l'évaluation. Il nous faut donc concevoir cet outil, c'est ce que nous réalisons dans le chapitre suivant.

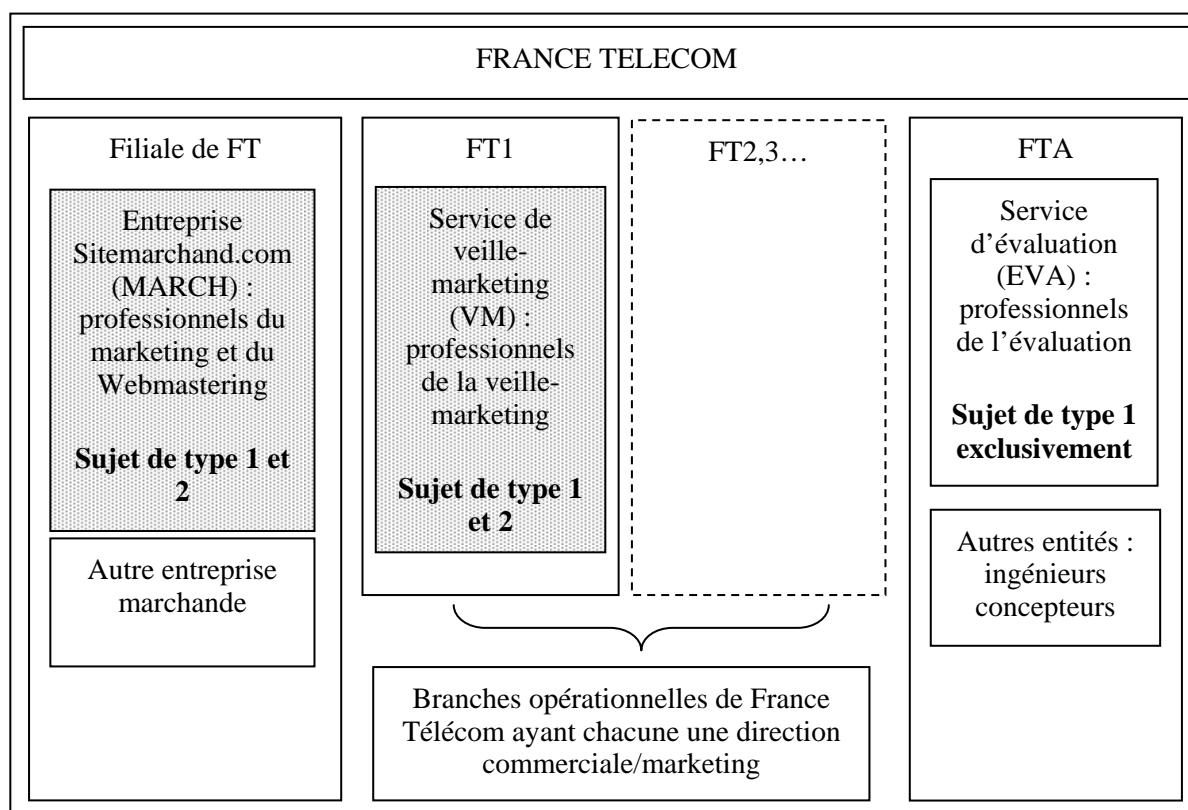
CHAPITRE 5

VERS LA CONCEPTION D'INSTRUMENTS POUR LES PROFESSIONNELS DU WEB MARCHAND

REPERES POUR LE CHAPITRE

Nous réutilisons le schéma de l'organigramme simplifié de France Télécom (cf. figure 16) pour resituer le travail proposé dans ce chapitre. Le centre d'intérêt n'est plus focalisé sur les professionnels du service d'évaluation (sujets de type 1 exclusivement) mais sur les professionnels du Web marchand (parties grisées dans le schéma). Nous avons vu que les professionnels du Web marchand ont leurs propres ressources d'évaluation de leur site marchand. Ils sont donc à la fois des sujets de type 1 (évaluateurs) et de type 2 (professionnels de la transformation). Dans ce chapitre, nous cherchons à étudier et à outiller leur activité.

Figure 16 : Organigramme simplifié des branches et filiales de France Télécom concernées dans cette thèse



Légende : Les sujets de type 1 ont une activité d'évaluation et les sujets de type 2 ont une activité de conception/marketing des produits/systèmes.

1. INTRODUCTION

Nous avons vu, au chapitre précédent, que les dimensions utilisateurs devaient être opérationnalisées pour les professionnels du marketing. La question qui se pose ici est de savoir quels outils il faut fournir aux professionnels du marketing, mais aussi aux professionnels de la conception, pour les aider dans leur propre activité d'évaluation. Il ne s'agit pas seulement de produire des résultats (même ergonomiques), puis de leur appliquer le format approprié pour les transmettre, mais véritablement de s'intéresser de plus près à l'activité de ces professionnels pour connaître ce qu'ils ont besoin de savoir pour agir et transformer leur produit/service. C'est ce que nous cherchons à faire, dans ce chapitre et le suivant, pour les professionnels du Web marchand (marketing et conception).

Nous cherchons à comprendre les exigences de leur activité et leurs représentations en vue de construire un outil d'aide à la conception de leurs propres évaluations, à partir du terrain donné, celui du service de veille-marketing et de l'entreprise Sitemarchand.com.

Notre objectif est de mieux connaître les objectifs de leur activité (prescrits ou non) et leurs contraintes : les caractéristiques des sites Internet sur lesquelles ils peuvent agir, les éléments relatifs à toute l'activité de commerce électronique sur lesquels ils doivent prendre des décisions en termes de modification, etc. Par ailleurs, quelles sont leurs connaissances en termes de techniques d'évaluation ? Bref, quels sont leurs besoins en termes de connaissances et d'outillage ?

Nous bâtissons d'abord un cahier des charges des futures caractéristiques structurelles et fonctionnelles de l'outil d'aide à l'élaboration d'évaluations en tenant compte des exigences de l'activité des professionnels du Web marchand, à partir essentiellement des données recueillies lors des entretiens avec les professionnels de la veille-marketing (décrits au chapitre précédent).

Nous identifions ensuite les besoins en connaissances des professionnels du Web marchand à partir des précédents entretiens ainsi que de celui du responsable du Webmastering (également décrit au chapitre précédent), d'un outil d'aide à l'analyse experte de sites de commerce (M.A.S.C.) et enfin d'un questionnaire d'une précédente évaluation du Sitemarchand.com. De plus, nous consolidons ces données par nos propres observations de plusieurs sites marchands et par des connaissances portant, tant sur les activités technico-commerciales relatives au commerce électronique, que sur les usages et la perception des cyber-consommateurs, issues de la littérature du domaine et de l'étude décrite au chapitre 3

(les deux techniques comparées sont ici appelées « entretiens utilisateurs »). Nous en utilisons cette fois-ci le contenu qui porte sur la perception des utilisateurs/clients et non plus les éléments de comparaison de l'entretien individuel et du *focus group*. Nous y faisons référence quand nécessaire, sans exemplifier l'ensemble du contenu, dans ce chapitre et dans le suivant.

Enfin, nous construisons un nouveau découpage correspondant à 13 catégories en rapport avec l'activité des professionnels du Web marchand, afin de faciliter les processus d'instrumentation et d'instrumentalisation de l'outil. Les catégories rassemblent, chacune, plusieurs besoins en connaissances spécifiques.

Ce travail de spécification est préalable à l'élaboration d'une ébauche de l'outil ou plutôt de la boîte à outils d'aide à la conception d'évaluations qui sera soumise aux jugements de professionnels de la veille-marketing et du Sitemarchand.com. Nous exposons les éléments de cette validation et les modifications opérées sur l'ébauche de la boîte à outils au chapitre suivant. La connaissance de l'activité des divers professionnels y est aussi approfondie.

Pour le découpage des 13 catégories comme pour l'élaboration de la boîte à outils, notre démarche est celle de l'engineering. Les choix sont opérationnels. Le découpage des 13 catégories ne suit pas une logique systématique, mais opérationnelle. Comme nous l'avons vu et le verrons plus en détail dans ce chapitre, l'un des professionnels de la veille-marketing catégorise les dimensions utilisateurs en suivant une logique « des artefacts » (l'accès, l'offre, le panier, etc.) et l'autre en suivant une logique des actions face au dispositif (« *Le client arrive sur le site* », « *Le client est sur le site* », etc.). Les critères de notre découpage ne correspondent pas à ces logiques systématiques. A partir de notre connaissance de l'activité des professionnels, nous cherchons à établir un découpage fonctionnel qui corresponde le plus possible aux représentations pour l'action (Weill-Fassina *et al.*, 1993) des professionnels du marketing et de la conception. L'objectif est qu'ils s'approprient l'outil aisément, dans leur activité d'élaboration d'évaluations. Cette démarche technologique consiste aussi à suivre le principe opérationnel suivant : élaborer une première ébauche, regarder si elle fonctionne et la modifier si nécessaire. Nous verrons, dans le chapitre suivant, que nous modifions l'ébauche qui suscite certaines critiques de la part des professionnels.

2. CAHIER DES CHARGES DE LA BOITE A OUTILS

Dans ce qui suit, nous présentons le cahier des charges de la boîte à outils que nous avons établi à partir essentiellement des données des précédents entretiens avec les professionnels de la veille-marketing.

2.1. L'OBJECTIF DE L'EVALUATION SOUS LA FORME D'UNE QUESTION

Nous avons vu, au chapitre précédent, que les dimensions utilisateurs sont très formelles et, de ce fait, trop abstraites pour les professionnels. A plusieurs reprises, ils emploient une question pour donner à voir l'objectif plausible de l'évaluation. Le premier professionnel de la veille-marketing est particulièrement prolix sur ce point : « *Donc on va essayer de savoir si les clients ont perçu qu'il existait bien des informations liées aux produits [...] Et puis il y l'autre axe, est-ce que les clients que je vise souhaitent ça ?* ».

D'ailleurs, une des dimensions experts des sites marchands (cf. chapitre 2) était exprimée, dans sa définition, sous la forme d'une question. Cela aidait beaucoup à sa compréhension. Un des professionnels de l'évaluation avait fait la remarque qu'une question était parfois plus explicite qu'un intitulé ou énoncé plus formel. Nous n'avons pas retenu ce point, mais finalement il nous paraît mériter une attention de premier ordre.

Ainsi pour chaque « besoin en évaluation » identifié, nous avons formulé une question simple qui explique l'objectif de l'évaluation : « les internautes ont-ils vu... ? sont-ils satisfaits... ? » (voir l'exemple dans le tableau 15 ci-après).

2.2. DECOUPAGE EN CATEGORIES GENERALES ET EN « BESOINS EN EVALUATION » PLUS SPECIFIQUES

Les professionnels de la veille-marketing ont souvent l'impression que des dimensions utilisateurs sont redondantes, comme si un thème a déjà été abordé dans une précédente dimension. Ou bien, au contraire, ils anticipent les choses, pensant qu'une dimension désigne des éléments qui ne sont traités qu'ultérieurement.

Pour éviter cela, nous avons formé des « grandes » catégories qui permettent d'instrumentaliser l'ensemble de la boîte à outils facilement : attribution d'une signification à l'outil (à quoi sert l'outil ?) à travers les intitulés de (moins d'une quinzaine de) catégories et

attribution d'une signification à chaque catégorie (quel est son contenu ? est-ce que la catégorie peut m'être utile dans mon projet d'évaluation ?).

Plusieurs des neuf thèmes de la transformation¹³¹ de surface de la liste de dimensions ont été conservés. Cependant, d'une logique en partie liée au parcours de l'utilisateur sur le site (« Avant l'arrivée sur le site », « Accueil », « Offre ou positionnement du site », « Communication et valorisation de l'offre », « Recherche », « Commande », etc.)¹³², nous sommes passé à une logique relative à l'utilisateur de la boîte à outils (quel intitulé lui permet de comprendre le contenu de chaque catégorie). Ces catégories ont été établies à partir des diverses données présentées au paragraphe 3 ci-après.

Il s'agit aussi de permettre de développer des schèmes d'utilisation appropriés. Pour que les professionnels puissent associer la boîte à outils à un schème de consultation rapide, nous avons listé les 13 catégories en préambule, puis chacune des catégories et ses « besoins en évaluation » spécifiques forment une nouvelle fiche avec un titre rendu bien visible. Chaque fiche a aussi été découpée en trois colonnes :

- la première contient les « besoins en évaluation » spécifiques : l'artefact matériel ou symbolique à évaluer (ex. la notoriété, l'illustration des produits, etc.) ou une de leur caractéristique (ex. l'utilité des informations-produits, la mise en valeur graphique des pages Internet, etc.),
- la seconde contient la question qui formule l'objectif de l'évaluation,
- enfin, la dernière, plus importante, contient tout le détail concernant ce qu'il s'agit précisément d'évaluer et comment l'évaluer.

Le principe est que l'utilisateur de la boîte à outils repère tout d'abord la catégorie qui l'intéresse, puis balaye rapidement les différents besoins en évaluation isolés dans la première colonne, les objectifs isolés dans la seconde colonne et décide ou non d'approfondir la lecture du contenu de l'évaluation.

Voici un extrait de la boîte à outils concernant la catégorie « Page d'accueil » et le besoin en évaluation « Compréhension immédiate de l'objectif du site » (tableau 15 ci-après).

Tableau 15 : extrait de l'outil d'aide à la conception d'évaluations de sites Internet marchands concernant la catégorie « Page d'accueil » et le besoin en évaluation « Compréhension immédiate de l'objectif du site »

Page d'accueil

Besoins en évaluation	Objectif de l'évaluation	Types et détails de l'évaluation
Compréhension immédiate de l'objectif du site	<p>Les internautes comprennent-ils, en un clin d'œil, ce qu'ils peuvent faire ou acheter sur le site ?</p> <p>Savent-ils, dès la page d'accueil, si le site correspond à leurs attentes ou si le site s'adresse à une autre cible clients ?</p>	<p>Etude par Q ou plus approfondie par FG ou E individuels avec utilisation du site, auprès de nouveaux visiteurs ou clients d'un site marchand (ou anciens visiteurs ou clients, s'ils se souviennent de leur première impression).</p> <ul style="list-style-type: none"> • Appréciation de leur compréhension immédiate de ce qu'ils peuvent faire sur le site (question oui/non), • Et/ou recueil de leurs diverses impressions générales (questionnement ouvert), • Appréciation de leur impression quant à l'adéquation du site par rapport à leurs besoins (est-ce que vous avez tout de suite eu le sentiment que le site vous concernait ?) <p>EA avec recueil de leur représentation quant à la cible à qui s'adresse la page d'accueil (ex. E individuel : identification des éléments sur la page d'accueil qui leur permettent de déduire à qui s'adresse le site ; ou FG avec test projectif : selon eux, la gamme de produits, les marques, le niveau des prix, l'ambiance générale de la page d'accueil correspondent à quel type de consommateur ?)</p> <p>Liste à titre indicatif :</p> <ul style="list-style-type: none"> • fonction de recherche par prix (sur certains sites – ex. de vin – cette fonction est perçue comme s'adressant à des personnes ayant des petits revenus), • présence de liens (ex. sites de vin) qui proposent une recherche de produits pour les non-initiés ou au contraire pour les connaisseurs, etc.
Réassurance
Rubriques & services principaux

2.3. PRISE EN COMPTE DES CARACTERISTIQUES DE DIFFERENTS COMMERCES ELECTRONIQUES

De plus, la boîte à outils vise à être un outil d'aide à l'évaluation, non pas spécifique au Sitemarchand.com, mais générique, permettant de traiter l'évaluation de tout commerce électronique, quelles que soient ses caractéristiques : son positionnement (ex. produits « discount » ou produits haut de gamme), son envergure commerciale (vente à l'international ou non), sa capacité de stockage et de livraison (par La Poste, par un transporteur interne), les fonctionnalités du système développées (ex. recherche par fourchette de prix ou par marques des produits), etc. Nous sommes donc parti des besoins en évaluation identifiés à partir de l'entreprise Sitemarchand.com, mais avons extrapolé aussi à tout autre site marchand.

Ainsi, pour chaque « besoin en évaluation », nous avons précisé quel type de commerce électronique est concerné selon ses caractéristiques et, si nécessaire, quelle population d'internautes il faut interroger (ex. clients étrangers).

De plus, nous avons listé différents éléments de l'interface (ex. logo, liens, icônes, etc.) ou relatifs au commerce (ex. prix, livraison, S.A.V.), susceptibles d'intervenir dans la perception des cyber-consommateurs, quand nécessaire. Ces listes doivent aider le professionnel à mieux cerner l'évaluation utile à son site marchand et à l'orienter sur des pistes d'amélioration, si le résultat s'avère mauvais (cf. exemple dans le tableau 15).

2.4. DEFINITION DU FORMAT DE L'ÉVALUATION

Nous nous sommes limité au cadre d'une évaluation par questionnaire (parfois appelée enquête de satisfaction) auprès de cyber-consommateurs d'un ou de plusieurs segments de la cible clients d'un site marchand. Ce genre d'évaluation est assez généraliste et peut porter sur différents thèmes à la fois. Nous avons suggéré (dans la troisième colonne) quel type de question il est possible de formuler pour le futur questionnaire (ex. fermée oui/non, sous forme d'échelles de Likert en 5 points ou une question ouverte ; cf. tableau 15).

Nous avons aussi précisé quand peut être utile une étude par entretiens individuels avec éventuellement une séquence d'utilisation du site marchand ou une étude par *focus groups*. Il s'agit alors d'études plus approfondies, en substitut ou en complément de l'évaluation par questionnaire, selon les possibilités ou volontés des acteurs du domaine.

Enfin, nous avons établi des besoins en évaluation très « pointus » qui concernent uniquement certains commerces électroniques et certaines situations (ex. évaluation de la simplicité de mémorisation de l'adresse URL uniquement lors du lancement d'un site). Ces besoins en évaluation requièrent des études très approfondies, supposent d'avoir lieu à un moment précis du développement du site marchand et/ou supposent une remise en question assez vaste de l'organisation ou des investissements de l'entreprise marchande.

A titre indicatif, nous précisons encore que l'outil vise à traiter des évaluations d'un commerce électronique existant, avec une cible clients déjà définie, c'est-à-dire des visiteurs/acheteurs identifiés (pour constituer des échantillons de population à étudier). Il ne s'agit pas de réaliser des études de marché, amont au déploiement d'un commerce, même si nous abordons des aspects marketing, puisqu'il s'agit de besoins en évaluation exprimés par ces professionnels.

3. DECOUPAGE DES 13 CATEGORIES ET IDENTIFICATION DES BESOINS EN EVALUATION

Nous présentons d'abord les données ayant servi au découpage des 13 catégories et à l'identification des besoins en évaluation, puis explicitons le résultat de ces opérations.

3.1. LES DONNEES

3.1.1. Une première ébauche des besoins en évaluation des professionnels de la veille-marketing

Au cours de premiers entretiens libres, décrits au chapitre précédent, nous avons recueilli une liste sommaire de besoins en évaluation du premier professionnel de la veille-marketing et les regroupements des dimensions du second professionnel. Nous avons réalisé une synthèse de ces éléments qui portent donc, en substance, sur la satisfaction des internautes quant :

- au choix de produits et leurs prix (appelé « l'offre » en marketing),
- aux conditions de mise en vente (modalités de livraison, facilités de paiement, etc.),
- aux informations sur les produits,
- à l'aide et à la facilité pour trouver des produits, puis pour les commander,
- aux aspects « Netmarketing » comme l'envoi d'offres ou d'informations plus ou moins personnalisées,
- et enfin, une dimension à part : la notoriété du commerce électronique qui est un élément majeur influençant le comportement des cyber-consommateurs.

3.1.2. Les regroupements des dimensions utilisateurs par les professionnels de la veille-marketing

Nous avons également repris les regroupements des dimensions utilisateurs réalisés par les professionnels de la veille-marketing et déjà décrits au chapitre précédent. Nous les rappelons pour mémoire et en présentons une interprétation de leur contenu pour le découpage des catégories de la boîte à outils.

3.1.2.1. Les regroupements du premier professionnel de la veille-marketing

Le professionnel de la veille-marketing, spécialiste des sites marchands, forme les regroupements ci-dessous, en suivant une logique liée aux artefacts matériels ou symboliques des sites Internet ou de l'activité marchande¹³³ :

- Notoriété,
- Facilité d'accès,
- Accueil,
- Offre,
- Communication et visibilité des " choses " & Informations produits,
- Facilité à trouver les produits,
- Panier & Commande,
- Sécurité,
- S.A.V.,
- Incontournables.

Comme nous l'avons développé au chapitre précédent, ce professionnel demande de situer les dimensions utilisateurs dans un contexte d'évaluation concret. Ses catégories sont ainsi en rapport avec des objectifs d'évaluation et parfois des objectifs de transformation des sites Internet marchands, même s'il ne l'explique pas toujours.

Par exemple, la notoriété constitue une catégorie à part, car selon ce professionnel, « *Satisfaction et notoriété, j'ai l'impression, que ça ne va pas de paire. On n'a pas, on n'est pas satisfait de [la notoriété] ..., on a une impression de la notoriété ou une connaissance ou une appréciation de la marque mais...* ». Autrement dit, soit l'objectif d'évaluation est de mesurer la connaissance de l'internaute quant à l'enseigne du site marchand, soit il s'agit de connaître sa satisfaction concernant les différents aspects du site marchand (l'offre commerciale, l'interface du site, etc.).

De même, concernant les catégories « Offre » et « Communication et visibilité des " choses " & Informations produits », le professionnel explique que :

- soit il s'agit de remettre en question seulement le choix de produits sélectionnés (ou encore le positionnement) par le service commercial,
- soit il s'agit de remettre en question la présentation des informations, prise en charge par le service Webmastering et aussi marketing (spécification d'une ambiance).

Le découpage des autres catégories n'est pas explicitement justifié par le professionnel, mais ces dernières peuvent être interprétées en relation aux différentes activités développées dans le commerce électronique :

- « Facilité à trouver les produits » concerne la conception de l'interface et la programmation des requêtes de recherche dans la base de données-produits du site marchand (recherche par mots-clefs). Les résultats de cette catégorie engendrent donc une éventuelle transformation assurée par les concepteurs du Webmastering et du service informatique.
- « Panier & Commande » concerne, de même, la conception de l'interface et donc les mêmes acteurs que précédemment, mais aussi les professionnels du marketing qui sont particulièrement soucieux et s'impliquent davantage dans la conception et l'évaluation de ce processus d'achat en ligne (d'après l'entretien avec ce professionnel et ceux présentés au chapitre suivant).
- « Sécurité » : les résultats de cette catégorie engendrent une éventuelle transformation assurée par les informaticiens, après prise de décisions par le responsable du Webmastering et la direction de l'entreprise quant aux choix technologiques ou stratégiques.
- « S.A.V. » : les résultats de cette catégorie engendrent une éventuelle amélioration qui concernent le service logistique et le service clients.
- « Incontournables » comprend des dimensions génériques relatives à la conception de l'interface (esthétique, rapidité de chargement, etc.). Les résultats de cette catégorie impactent essentiellement l'activité des concepteurs du Webmastering et du service informatique, mais aussi ceux du marketing (vocabulaire, ambiance à définir).

3.1.2.2. Les regroupements du second professionnel de la veille-marketing

La logique retenue par le second professionnel de la veille-marketing, non spécialiste des sites marchands, est celle des actions face au dispositif¹³⁴ : *"Le client arrive sur le site"*, *"Le client est sur le site"* jusqu'à *"Le client reçoit le produit"*.

La discussion avec ce professionnel nous apporte moins de réponses concernant ses besoins en connaissances. Il n'a pas l'expérience de la construction d'un questionnaire de satisfaction sur les sites marchands, contrairement au précédent. Il se projette plus difficilement dans une situation d'évaluation d'un site marchand où il s'agit de se demander quels résultats seraient utiles. Il est ainsi plus centré sur le regroupement des dimensions sans

songer à expliciter ou sans pouvoir expliciter ce qui, par ailleurs, est évident pour lui, c'est-à-dire ce qu'il ne s'est jamais formulé lui-même (du moins de la façon dont nous le lui demandons). Il éprouve donc des difficultés à le faire : « *Oui c'est important... C'est pas évident ; ça me semble tellement évident, tu vois...* ». Nous nous retrouvons là face à la même problématique que celle des sujets naïfs qui répondent à une question d'évaluation à laquelle ils n'ont peut-être jamais réfléchi et où la question elle-même structure une opinion qu'éventuellement ils ne se forgent qu'à cette occasion (Ghiglione & Matalon, 1998).

3.1.3. L'outil d'aide à l'analyse experte de sites de commerce (M.A.S.C.)

Nous avons aussi complété les précédentes données par l'outil M.A.S.C. (méthode d'analyse de sites de commerce). Cet outil a été créé par un ergonome de FTA, impliqué dans des activités de conception de sites marchands. Il permet l'analyse comparative (« benchmark ») de sites de commerce électronique par un professionnel du marketing ou chef de projet selon huit fondamentaux, comportant chacun une série d'éléments définis en termes de ce qui est attendu d'un « bon » site marchand (ex. le site doit être mentionné dans les cinq principaux moteurs de recherche, ce que l'entreprise marchande doit présenter comme garanties de livraison, etc.).

Cette méthode nous semble intéressante car elle cherche à outiller les professionnels du Web marchand. De plus, son élaboration repose sur des données relatives à l'activité de ces professionnels. La construction de ses huit fondamentaux et de leurs éléments constitutifs est, en effet, basée sur les connaissances opérationnelles de l'ergonome en interaction avec des professionnels du Web marchand (complétées de données issues de cinq *focus groups* sur les freins et facteurs d'achats des cyber-consommateurs). Les fondamentaux et éléments constitutifs sont aussi testés à partir de l'analyse comparative de 20 sites marchands par deux paires d'étudiants d'une école de commerce.

Nous nous sommes inspiré de cet outil mais avons poursuivi notre logique qui est d'identifier les besoins en évaluation du Web marchand, non pas pour une expertise comparative de sites, mais pour une évaluation auprès d'internautes. Les questions ne se posent, en effet, pas dans les mêmes termes.

Prenons un exemple concret : dans M.A.S.C., chaque élément de chacun des huit fondamentaux est noté par l'analyste en fonction de sa qualité ou de sa présence/non

présence, comme l'élément « livraison avec rendez-vous (avec une plage horaire précise) ». Or, lors d'une enquête de satisfaction d'un site marchand, il s'agit de considérer les caractéristiques propres à chaque entreprise marchande : la plupart de celles proposant la prestation de livraison avec rendez-vous sont des cyber-commerces de comestibles (fleurs, petits-déjeuners à domicile, etc.), tandis que de très nombreux sites marchands de produits non comestibles et peu encombrants (produits culturels, jouets, petit électroménager, etc.) choisissent La Poste comme partenaire d'expédition. Les besoins en évaluation de chacun sont donc différents. Pour la plupart des sites marchands utilisant les services de La Poste, nul n'est besoin de remettre en question ce choix et d'opter pour un transporteur, dont les services seraient beaucoup plus chers. Une évaluation subjective de l'utilité d'une livraison par rendez-vous n'est alors pas justifiée.

3.1.4. L'entretien avec le responsable du Webmastering

Nous avons aussi réutilisé l'entretien de validation de la transformation de la liste de dimensions, avec le responsable du Webmastering, décrit au chapitre précédent, qui nous apporte des connaissances concernant son activité. De plus, nous apprenons, en partie, le rôle des professionnels du marketing, du service produits, du service logistique et du service clients. Pour plus d'efficacité, nous ne présentons pas tous ces éléments, mais les réutilisons, au fur et à mesure, quand nécessaire, dans ce chapitre et dans le suivant.

3.1.5. Une trace de l'activité : un questionnaire d'évaluation du Sitemarchand.com

Enfin, nous avons exploité un questionnaire d'une enquête de satisfaction, réalisée après la période de Noël 2001, pour mesurer l'appréciation subjective de l'entreprise Sitemarchand.com dans son ensemble, de son offre commerciale, de l'efficacité de son service clients, de ses performances logistiques et de conception du système (rapidité d'accès, facilité du processus d'achat, etc.).

Le professionnel de la veille-marketing, spécialiste des sites marchands, a contribué à l'élaboration de cette étude dont les besoins en évaluation ont été identifiés par l'ensemble des responsables des différents services (produits, logistique, clients, marketing, Webmastering) de l'entreprise Sitemarchand.com et synthétisés par le professionnel du marketing, chargé des études. Il s'agit d'une étude contextualisée dépendant de besoins en

évaluation ponctuels, liés aux diverses caractéristiques du Sitemarchand.com au moment de l'évaluation (après Noël), à la maturité de l'activité marchande (phase de fidélisation de clients déjà acquis) et aux connaissances des professionnels relatives aux caractéristiques de leur clientèle. Ainsi, les besoins en évaluation que nous avons identifiés à partir de ce questionnaire ne constituent pas une formalisation complète. Néanmoins ce questionnaire nous apporte une indication forte.

3.2. LE DECOUPAGE DES 13 CATEGORIES ET LEUR CONTENU

Dans le découpage proposé ici (cf. tableau 16 ci-après), notre objectif est de s'approcher le plus possible des catégories fonctionnelles (Dubois, 1997) des professionnels du marketing et du Webmastering, donc de leurs représentations pour l'action. Il s'agit de faire en sorte que l'intitulé de chaque catégorie évoque le plus possible ce qu'elle contient au plus grand nombre de professionnels. Ce découpage doit, en effet, aider les professionnels à s'approprier la boîte à outils dans leur activité d'élaboration d'évaluations. Il n'a pas « valeur de vérité » ; il peut d'ailleurs être plus ou moins modifié par chaque professionnel, selon son domaine d'activité et ses contraintes spécifiques.

Nous nous sommes davantage appuyé sur le découpage du premier professionnel de la veille-marketing car, en dehors de sa logique basée sur les artefacts, il peut être interprété en rapport avec les différentes activités du Web marchand. Nous avons poursuivi et accentué cette dernière logique. Ces catégories incorporent également plusieurs des huit fondamentaux de l'outil M.A.S.C.¹³⁵ puisque cet outil est élaboré à partir d'un matériel issu des professionnels du Web marchand et destiné à être utilisé par eux. Le second professionnel de la veille-marketing, lors de son premier entretien, nous fait d'ailleurs part de son approbation quant à cet outil : il l'a mémorisé et peut ainsi facilement réutiliser les fondamentaux de cet outil. Les huit fondamentaux sont néanmoins insuffisants à contenir tous les besoins en connaissances identifiés ici et les évaluations possibles proposées dans notre boîte à outils.

Tableau 16 : 13 catégories correspondant aux besoins en connaissances des professionnels du marketing et de la conception dans l'évaluation des sites Internet marchands

<ul style="list-style-type: none">• Notoriété• Facilité d'accès & Capacité du système• Page d'accueil • Offre & Positionnement• Description des produits• Look, présentation, visibilité et navigation• Netmarketing • Sécurité & Paiement• Interactivité• Logistique • Les informations associées au service de vente • Facilité à trouver un produit• Facilité à commander un produit

Concernant le contenu des catégories, en plus des données empiriques, nous nous sommes appuyé sur de nombreux travaux qui corroborent les besoins en connaissances exprimés par les professionnels ou leurs connaissances des attitudes et comportements de l'utilisateur/client (qui concourent à l'existence d'un besoin en évaluation).

Nous développons, dans ce qui suit, chacun des besoins en connaissances identifiés dans les 13 catégories.

3.2.1. Notoriété

Les deux professionnels de la veille-marketing s'entendent pour dire que la notion de notoriété est un élément fondamental qui va influencer le comportement du consommateur de façon déterminante. Le premier professionnel de la veille-marketing affirme par exemple : « *Parce que la notoriété c'est la notoriété globale de l'entreprise au bout du compte* », « *J'aurais tendance à dire que notoriété c'est un élément d'ensemble qu'on va apprécier sans même avoir mis le nez sur l'écran et qui va évoluer après, en fonction de ce qu'on perçoit* » et le second professionnel : « *Bon la notoriété, à mon avis, induit beaucoup de choses* ». En effet, il est démontré que la notoriété de l'entreprise marchande a un impact significatif sur la

confiance accordée à un commerce électronique et sur la facilité à transmettre ses coordonnées bancaires (Jarvenpaa *et al.*, 2000; Kehoe *et al.*, 1998a; Nielsen *et al.*, 2001).

De même, l'image de marque (Dayan, 2001), notion complémentaire à la notoriété, peut avoir un effet de « halo » et influencer la perception de la qualité de tout un commerce électronique (ses produits, son service après vente, etc.; Dayan, 2001; Lohse & Spiller, 1998a). Le professionnel de la veille-marketing spécialiste des sites marchands tient aussi compte de cet élément : *« ce qu'on peut se poser comme question c'est si le site lui-même, si la marque, si l'effet qu'il donne est positif ou négatif par rapport à la notoriété globale de l'entreprise [...] Mais il l'associera aussi [un site marchand] à la qualité de ses livraisons, aux prix de ses produits, au choix.... »*.

Tout cela nous a permis de constituer les besoins en évaluation suivants (cf. aussi tableau 17, page suivante) :

- la connaissance des clients concernant l'enseigne (la notoriété ou réputation de la marque, de l'entreprise marchande « terrestre ou virtuelle »).
- l'image de marque de l'entreprise marchande liée à la satisfaction des internautes concernant la qualité des produits, l'efficacité du service après vente, etc.

Le questionnaire d'évaluation de Noël 2001 et Macarez et Leslé (2001) nous ont aussi aidé à discerner deux autres besoins en évaluation :

- l'impact de la médiatisation de l'entreprise marchande sur la découverte de son site : il s'agit d'estimer la proportion d'internautes de la cible clients d'un site marchand l'ayant découvert par les différents supports de médiatisation ou par le bouche à oreille,
- les facteurs de retour des visiteurs sur le site (une publicité, un bon d'achat, par satisfaction des précédents achats, etc.).

Enfin, un développement sur la « mesure de l'efficacité publicitaire » de Dayan (2001) nous a conduit à proposer des études plus poussées qui doivent permettre de recueillir des informations concernant :

- l'efficacité de la médiatisation « offline » de l'entreprise sur ses ventes et sa notoriété,
- et la compréhension des messages publicitaires.

- Récapitulatif des différents besoins en évaluation dans chaque catégorie et les sources ayant permis de les identifier :

Le tableau 17 ci-dessous et les tableaux suivants (18 à 29) récapitulent les 13 catégories et les différents besoins en évaluation qu'elles contiennent. Il est également précisé les éléments qui nous ont conduit à les former. Il s'agit :

- des données des entretiens des professionnels de la veille-marketing (Entretiens prof. VM.),
- de l'entretien avec le professionnel du Webmastering (Entretien resp. Webmastering),
- de l'outil M.A.S.C.,
- du questionnaire d'évaluation de Noël 2001 (Questionnaire),
- ou enfin de connaissances issues de la littérature du domaine ou de documentations techniques nous permettant de mieux connaître les activités relatives au marketing du commerce électronique et traditionnel, à la logistique de la vente par correspondance, à la sécurisation des diverses données personnelles par Internet, au service clients, à la conception de pages Internet, etc. (Littérature/Documentation).

Tableau 17 : Besoins en connaissances des professionnels du Web marchand concernant la Notoriété

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Notoriété	Entretiens prof. VM. + Outil M.A.S.C.+ Littérature/Documentation
Image de marque	Entretiens prof. VM + Littérature/Documentation + Questionnaire
Efficacité de la médiatisation par rapport à la découverte du site	Questionnaire + Littérature/Documentation + Outil M.A.S.C.
Médiatisation, image de marque et raisons de retour des visiteurs sur le site	Questionnaire
Efficacité de la médiatisation « offline » du site sur ses ventes et sa notoriété	Littérature/Documentation
Réception des messages publicitaires	Littérature/Documentation

3.2.2. Facilité d'accès & Capacité du système

Le premier professionnel de la veille-marketing forme un regroupement « Facilité d'accès » avec le référencement du site sur les outils de recherche sur le Web et la simplicité de l'adresse URL (ex. « *Cdiscount, on ne sait pas s'il y a un ou deux "d". Ce sont des détails mais...* »). Selon lui, ces deux aspects sont « *des choses un peu techniques, techniques d'Internet* ».

Par ailleurs, on sait que la première page des quinze principaux outils de recherche oriente 95% des recherches sur Internet. Ainsi, « *une campagne de publicité ne peut remplacer un référencement efficace* » (Macarez & Leslé, 2001). Il s'agit donc, de même que pour la médiatisation du site, de savoir quel rôle le référencement joue sur la découverte du site et sur sa fréquentation. Nous avons également proposé d'évaluer la simplicité de mémorisation de l'adresse URL du site Internet.

Le premier professionnel de la veille-marketing a aussi spontanément cité « *le poids de la page d'accueil* », c'est-à-dire l'évaluation de la satisfaction concernant la rapidité de chargement du site, comme faisant partie de cet ensemble « technique ». Le besoin d'évaluer ce point est conforté par les nombreuses études montrant que la lenteur de chargement constitue une frustration majeure pour les internautes (Katz & Aspden, 1997; Kehoe *et al.*, 1998a; Manchester_Institute_for_Popular_Culture, 1999; Pitkow & Kehoe, 1996).

Tableau 18 : Besoins en connaissances des professionnels du Web marchand concernant la Facilité d'accès et la Capacité du système

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Référencement	Littérature + Entretien resp. Webmastering + Outil M.A.S.C. + Littérature/Documentation
Simplicité Adresse URL	Entretiens prof. VM + Entretiens utilisateurs
Rapidité de chargement	Entretiens prof. VM + Littérature/Documentation + Questionnaire

3.2.3. Page d'accueil

En plus des craintes relatives au paiement en ligne, la facilité des internautes à « zapper » d'un site à un autre est bien connue dans le domaine (« one click away » ; Lindgaard & Dudek, 2002; Macarez & Leslé, 2001). La page d'accueil doit ainsi être une véritable carte de visite ou « vitrine » (Fouchard, 2001) et assurer un certain nombre de fonctions essentielles. Le premier professionnel de la veille-marketing acquiesce : « *C'est vachement important en effet ce que tu disais là. Comprendre clairement de quoi il s'agit. Qu'est-ce qu'on y fait sur ce site ? [...] c'est savoir ce qu'on y fait, ce qu'on y vend [...] C'est comme une vitrine, en fait. Quand tu vas sur le site, on voit ce qu'on y vend. Quelle est la gamme de produits vendus ? Quel est le positionnement du magasin ? On voit tout de suite si c'est pour nous ou si c'est pour quelqu'un d'autre* ».

En plus des propos des professionnels, les divers besoins en connaissances et la catégorie elle-même nous ont été inspirés par nos entretiens utilisateurs et divers travaux (cf. tableau 19 ci-après).

Cette catégorie cherche à vérifier si les fonctions suivantes sont remplies par la page d'accueil :

- la compréhension immédiate de ce que l'internaute peut faire ou acheter sur le site et se reconnaît-il dans la cible clients à qui le site cherche à s'adresser (Fouchard, 2001; Nielsen, 2000; Nielsen *et al.*, 2001),
- la fonction de réassurance, relative à la présence d'informations sur l'entreprise (ex. coordonnées, logo de partenaires financiers ou d'organismes de contrôle, etc.) (Fouchard, 2001; Pilverdier & Gaillard, 2001),
- l'accès facile aux rubriques et aux services principaux du site (Fouchard, 2001; Leulier *et al.*, 1998 ; Nielsen, 2000; Nielsen *et al.*, 2001; Nogier, 2002).

Tableau 19 : Besoins en connaissances des professionnels du Web marchand concernant la Page d'accueil

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Compréhension immédiate de l'objectif du site	Entretiens prof. VM + Littérature/Documentation
Réassurance	Entretiens utilisateurs + Entretiens prof. VM + Littérature/Documentation + Outil M.A.S.C.
Accès à tous les services et rubriques	Entretiens utilisateurs + Littérature/Documentation

3.2.4. Offre & Positionnement - Description des produits – Look & co

La distinction de ces trois catégories nous a été dictée par les propos du premier professionnel de la veille-marketing qui demande de distinguer s'il s'agit de remettre en question l'offre et le positionnement du commerce électronique ou bien la présentation de cette offre à l'interface.

Dans le regroupement « Communication/visibilité des " choses " & Informations produits » du professionnel, on peut encore logiquement distinguer s'il s'agit de remettre en question le contenu des informations décrivant les produits ou les aspects visuels de la présentation de ces informations. Pour ce découpage, nous nous sommes aussi appuyé sur celui observé par plusieurs auteurs (Kim & Lee, 2002; Kwon *et al.*, 2002; Shneiderman, 1998) qui distinguent l'architecture ou la structure d'une interface, son contenu informatif et la présentation de l'information. De même, le processus de conception des sites Internet suit trois étapes majeures similaires : l'identification des contenus selon l'objectif de l'entreprise et les besoins des futurs utilisateurs, leur organisation architecturale ainsi que les principes de navigation et enfin la conception des aspects graphiques de chaque page Internet (Goto & Cotler, 2002).

C'est ce qui nous a conduit à créer les catégories « Description des produits », d'une part, et « Look, présentation, visibilité et navigation » (ou aussi « Look & co »), d'autre part. Cette dernière catégorie comprend des aspects visuels et quelques aspects architecturaux et « navigationnels » (voir paragraphe 3.2.7.).

Ces trois catégories différentes correspondent aussi à des actions d'amélioration et des acteurs différents :

- si c'est un problème de définition de l'offre, la responsabilité revient aux professionnels du marketing : il faut réviser les orientations choisies face à la concurrence et réaliser des études approfondies pour mieux cerner les attentes de la cible clients,
- s'il faut améliorer les éléments de contenu qui décrivent les produits : le service produits en charge de recueillir les informations-produits (éventuellement le service éditorial, s'il y en a un) sera sollicité pour trouver de meilleurs éléments d'information,
- enfin, s'il faut seulement revoir la conception de l'interface : l'équipe marketing va redéfinir des positionnements identitaires et de nouvelles ambiances, et les Webdesigners les matérialiseront (nouvelles pages Internet).

3.2.5. Offre & Positionnement

Plusieurs travaux mettent en évidence des corrélations significatives entre la taille perçue de l'offre commerciale, d'une part, et la confiance accordée au commerce électronique et l'intention d'acheter sur le site, d'autre part (Jarvenpaa *et al.*, 2000; Konradt *et al.*, 2003). Ces études et les propos des deux professionnels de la veille-marketing attestent qu'il est nécessaire d'évaluer la perception de l'offre de produits.

D'après l'ensemble des données empiriques recueillies (présentées plus haut) et plusieurs travaux du domaine, il nous semble qu'un ensemble d'évaluations pourrait porter sur la satisfaction des cyber-consommateurs concernant (cf. aussi tableau 20 ci-après) :

- le choix de produits/services et notamment la variété ou la profondeur de l'offre (Jarvenpaa & Todd, 1997) :
 - la variété ou encore taille perçue : nombre de familles de produits différents (CD, hifi, livres, etc.),
 - la profondeur : étendue du choix dans chaque famille, gammes de produits,

- la qualité des produits (importance de l'information permettant de décider de la qualité d'un produit ; Kehoe *et al.*, 1998b; Kim & Lee, 2002; Lee *et al.*, 2000; Liu & Arnett, 2000; Manchester_Institute_for_Popular_Culture, 1999),
- le choix de marques proposées (Jarvenpaa & Todd, 1997; Nielsen *et al.*, 2001),
- l'intérêt/quantité de produits/services rares (Jarvenpaa & Todd, 1997; Nielsen *et al.*, 2001),
- le niveau des prix (Jarvenpaa & Todd, 1997),
- le choix de promotions et de nouveautés (Fouchard, 2001; Jarvenpaa & Todd, 1997; Lohse & Spiller, 1998b; Rehman, 2000).

Par ailleurs, le premier professionnel de la veille-marketing nous explique que le commercial doit être attentif à son positionnement face à la concurrence qu'il s'agisse d'autres sites Internet ou de magasins de proximité du même secteur de vente (que l'entreprise marchande ait des magasins traditionnels ou qu'elle soit un « pure player »¹³⁶). De même, Jarvenpaa et Todd (1997) formulent une recommandation en ce sens aux professionnels du Web marchand.

Dans le cas où l'entreprise marchande n'est pas un « pure player », le professionnel de la veille-marketing nous explique qu'il s'agit aussi d'adapter l'étendue et l'intérêt de l'offre en ligne, comparativement à celle de ses propres magasins « réels ». Nos entretiens utilisateurs et plusieurs travaux (Lohse & Spiller, 1998a; Nielsen *et al.*, 2001) montrent effectivement que certains internautes préfèrent le site marchand qui propose le même choix de produits et de promotions que dans son magasin « réel », plutôt que celui qui ne le fait pas.

Tout cela a donné naissance à deux autres axes de besoins en connaissances :

- le positionnement par rapport à la concurrence,
- et le positionnement interne.

Tableau 20 : Besoins en connaissances des professionnels du Web marchand concernant l'Offre et le Positionnement

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Choix de produits/services	Entretiens prof. VM + Questionnaire + Outil M.A.S.C. + Entretiens utilisateurs + Littérature/Documentation
Prix	Entretiens prof. VM + Questionnaire+ Outil M.A.S.C.+ Entretiens utilisateurs + Littérature/Documentation
Promotions & Nouveautés	Entretiens prof. VM+ Outil M.A.S.C. + Entretiens utilisateurs + Littérature/Documentation
Positionnement interne (même enseigne)	Entretiens prof. VM + Entretiens utilisateurs + Littérature/Documentation
Positionnement par rapport à la concurrence	Entretiens prof. VM + Littérature/Documentation

3.2.6. Description des produits

L'étude de Kim et Lee (2002) révèle que la phase de prise d'informations sur le produit (*Information phase*) est la plus importante, devant la phase de mise en panier (*Agreement phase*), devant la phase de « finalisation et réception de la commande » (*Settlement phase*) et devant une phase dite *Environment phase* qui concerne les éléments de l'interface intervenant aux trois phases précédentes. Une fois de plus, cela confirme ce que les professionnels de la veille-marketing nous ont déjà énoncé : cette catégorie constitue un besoin en évaluation.

Les divers besoins en évaluation de cette catégorie nous ont été suggérés par toutes les données empiriques et la littérature du domaine (cf. tableau 21 ci-après). Ils concernent la satisfaction des internautes par rapport :

- au choix des éléments de définition des produits (Kim & Lee, 2002), leur complétude (Jarvenpaa & Todd, 1997; Lee *et al.*, 2000) et leur clarté (Konradt *et al.*, 2003),
- aux commentaires ou critiques de professionnels/internautes sur les produits (Kim & Lee, 2002),
- aux illustrations (Pilverdier & Gaillard, 2001),
- à la façon d'annoncer les prix (Nielsen *et al.*, 2001; Pilverdier & Gaillard, 2001),
- à l'annonce de la disponibilité et des frais d'envoi (de chaque produit individuellement) (Fouchard, 2001; Nielsen *et al.*, 2001; Pilverdier & Gaillard, 2001),
- à la mise à jour de tous ces éléments et aux informations qui permettent de connaître la dernière date de mise à jour (Fouchard, 2001),
- à leur visibilité (cf. catégorie « Look & co »).

Enfin, une évaluation particulière a été ajoutée concernant l'éditorial car le premier professionnel de la veille-marketing pense que leur intérêt serait d'avoir une indication sur la

valeur ajoutée que représente un véritable éditorial (informations plus fouillées, rédigées par un vrai service éditorial qui présente donc un certain coût). La valeur ajoutée d'un cyber-commerce par rapport à un commerce traditionnel résiderait, en effet, dans sa capacité à prodiguer toutes sortes de conseils et d'apporter de l'information dans le domaine (Fouchard, 2001).

Tableau 21 : Besoins en connaissances des professionnels du Web marchand concernant la Description des produits

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Caractéristiques produits	Entretiens utilisateurs + Entretiens prof. VM + Outil M.A.S.C. + Littérature/Documentation
Commentaires & Critiques	Entretiens utilisateurs + Entretiens prof. VM + Outil M.A.S.C. + Littérature/Documentation
Illustrations	Entretiens utilisateurs + Outil M.A.S.C. + Littérature/Documentation
Annonce prix	Entretiens utilisateurs + Entretiens prof. VM + Littérature/Documentation
Annonce disponibilité & frais d'envoi	Entretiens utilisateurs + Questionnaire + Outil M.A.S.C. + Littérature/Documentation
Mise à jour	Entretiens prof. VM + Outil M.A.S.C.
Visibilité des informations-produits	Cf. Catégorie suivante « Look & co »
Editorial (valeur ajoutée)	Entretiens prof. VM

3.2.7. Look, présentation, visibilité et navigation

Concernant la catégorie « Look, présentation, visibilité et navigation » qui s'applique à toutes les pages d'un site marchand (y compris les courriels comme les « Newsletters » envoyés aux clients), nous avons rassemblé les aspects hédoniques (éléments attractifs, ludiques ou au contraire ringards, etc.) et les aspects graphiques qui peuvent aider la lecture et la navigation sur le site (visibilité des informations importantes, des boutons d'action, etc.).

L'activité de conception des pages Internet nous conduit à rassembler ces deux aspects. Quand un Webdesigner conçoit une page, il tient compte des dimensions « esthétique » et « efficacité visuelle ». Généralement, les professionnels du marketing du Sitemarchand.com spécifient l'ambiance recherchée (ex. il faut que ce soit « promo » et « fun » pour une offre spéciale sur un assortiment de produits culturels pour les jeunes). Le Webdesigner conçoit alors plusieurs concepts qui correspondent à cette attente, tout en respectant la définition de la charte graphique de l'entreprise (typologie de couleurs, de formes, de polices, etc.) (Goto & Cotler, 2002; Nogier, 2002). Des aspects ergonomiques sont aussi pris en compte de façon spontanée (Chevalier & Ivory, 2003). Puis, le responsable de la charte graphique ou le

responsable du Webmastering sélectionne un des concepts et vérifie (intuitivement) si les aspects « visibilité » et « navigation » sont acceptables.

A partir de l'entretien avec le responsable du Webmastering, de nos entretiens utilisateurs, de l'outil M.A.S.C. et de quelques travaux issus de la littérature (cf. tableau 22 ci-dessous), nous avons proposé l'inventaire d'évaluations de la satisfaction des internautes concernant :

- l'esthétique du site (ex. harmonie des couleurs) (Lindgaard & Dudek, 2002),
- l'attractivité du site, l'ambiance dégagée par le site (Hassenzahl & Trautmann, 2001; Leulier *et al.*, 1999; Liu & Arnett, 2000; Park *et al.*, 2004),
- la qualité des graphismes (Nogier, 2002),
- la facilité à reconnaître le site visuellement (quelques internautes de nos entretiens utilisateurs veulent pouvoir reconnaître assurément l'entreprise marchande, pour être sûrs de ne pas être sur un site pirate),
- l'aération générale des pages (Nielsen, 1996),
- l'organisation, la mise en page des informations et leur mise en valeur graphique (Leulier *et al.*, 1998; Nielsen, 2000; Nogier, 2002),
- la visibilité des informations *importantes* (ex. les adresses électroniques et numéros de téléphone des divers services, etc.),
- la visibilité et les emplacements des liens, barres de navigation et boutons d'action (Leulier *et al.*, 1998; Nielsen, 2000; Nielsen *et al.*, 2001; Nogier, 2002; Tilson *et al.*, 1998).

Tableau 22 : Besoins en connaissances des professionnels du Web marchand concernant le Look, la Présentation, la Visibilité et la Navigation

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Esthétique	Entretiens utilisateurs + Littérature/Documentation
Attractivité	Entretien resp. Webmastering + Entretiens utilisateurs+ Littérature/Documentation
Graphismes	Entretiens utilisateurs + Littérature/Documentation
Reconnaissance visuelle	Entretien resp. Webmastering + Entretiens utilisateurs
Aération	Entretien resp. Webmastering + Littérature/Documentation
Organisation, Mise en page, Mise en valeur graphique	Entretiens utilisateurs + Littérature/Documentation + Outil M.A.S.C.
Visibilité des informations importantes	Entretiens utilisateurs + Outil M.A.S.C. + Littérature/Documentation
Visibilité & Emplacements des boutons d'action	Entretiens utilisateurs + Littérature/Documentation + Outil M.A.S.C.

3.2.8. Netmarketing

Avec l'Internet, le commerce a changé de forme et le marketing *one-to-one* s'est développé (Macarez & Leslé, 2001). Les entreprises marchandes déploient un effort considérable pour apporter une information dite personnalisée à leurs clients, mais aussi pour promouvoir leurs produits sur le site directement. Les professionnels du Web marchand ont donc besoin de savoir si leurs diverses actions de promotion ont un impact sur leur clientèle.

L'outil MA.S.C. et divers travaux concernant le « marketing en ligne » (cf. tableau 23 ci-après) nous ont permis de dégager un ensemble de besoins en évaluation qui concerne :

- la qualité du contenu des courriels d'informations ou d'offres personnalisées et leur quantité (Fouchard, 2001; Lohse *et al.*, 2000; Macarez & Leslé, 2001),
- l'intérêt de la vente croisée (« Cross-Sellings »), c'est-à-dire des propositions d'accessoires ou de produits supplémentaires (Fouchard, 2001; Macarez & Leslé, 2001),
- l'intérêt du contenu des présélections dans les listes de résultats issues d'une recherche de produits (ex. les trois produits les plus vendus en début de liste),
- l'intérêt des avantages clients (réduction 1^{ier} achat, bon d'achat, etc.) (Lohse & Spiller, 1998a) et l'utilisation de cartes privatives,
- la qualité et la quantité de la publicité en ligne (Carton, 2002; Nielsen *et al.*, 2001; Spool *et al.*, 1999).

Nous avons placé les envois de « Newsletters » et la vente croisée dans cette catégorie « Netmarketing » même si le premier professionnel de la veille-marketing les place dans la catégorie « Offre », lors des regroupements qu'il a réalisés¹³⁷. En effet, ce professionnel nous fait aussi part de son activité indépendante d'analyse « benchmark » des « Newsletters » distribuées par le Sitemarchand.com et par les sites concurrents. Nous pensons donc que les professionnels du marketing ont besoin de savoir s'il s'agit de remettre en question l'offre générale de produits, d'une part, ou les offres croisées, personnalisées, etc., d'autre part. Il s'agit de deux activités complémentaires mais différentes. De plus, ces courriels ou offres, de même que la publicité en ligne, requièrent une activité de marketing ou de conception particulière : création de « Newsletters », paramétrage d'outils de recueil de données sur les internautes, exploitation de bases de données, identification de profils de clients, etc. Il nous semble ainsi logique de rassembler l'ensemble de ces éléments dans cette catégorie « Netmarketing » et ceci sera confirmé ou infirmé à partir des entretiens de validation de l'ébauche.

Tableau 23 : Besoins en connaissances des professionnels du Web marchand concernant le Netmarketing

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
« Cross-Sellings »	Entretiens prof. VM + Littérature/Documentation
Présélections dans les listes de résultats	Entretiens prof. VM + Littérature/Documentation
« Newsletters » non personnalisées	Entretiens prof. VM + Littérature/Documentation + Questionnaire
« Newsletters » ou offres personnalisées	Entretiens prof. VM + Littérature/Documentation + Questionnaire
Offres personnalisées (« Opt-in »)	Entretiens prof. VM + Littérature/Documentation
Courriels en « push » (« Opt-out »)	Entretiens prof. VM + Littérature/Documentation
Publicité en ligne	Entretiens utilisateurs + Littérature/Documentation
Avantages clients & Cartes privées	Entretiens prof. VM + Outil M.A.S.C.

3.2.9. Sécurité & Paiement

Concernant l'aspect sécurité, les professionnels le disent, c'est un élément important en matière de commerce électronique, néanmoins ils l'affirment aussi, il n'y a pas de latitude dans le choix des techniques de sécurisation : le serveur S.S.L. (*Secure Sockets Layer*) avec son cryptage des données bancaires. Le second professionnel de la veille-marketing tient, en effet, les propos suivants : « *Donc y a pas à tergiverser là dessus, ça l'est : donc le client il part dans ce contexte là, où il arrive sur quelque chose qui est balisé au niveau sécurité* », « *Par contre, je serais vigilant dans ma conception de service à qu'est-ce que... effectivement ça l'est dans la page d'accueil [indications sur la sécurité] et dans le processus [sécurité effective], ce serait une contrainte à gérer* ». Il s'agit donc d'une condition qui doit être remplie par le système, mais qui n'est pas suffisante pour attirer l'internaute (Liu & Arnett, 2000).

Néanmoins, comme dans le questionnaire Noël 2001, et sur la base de nos entretiens utilisateurs et de plusieurs travaux (cf. tableau 24 ci-après), nous avons proposé d'évaluer trois éléments relatifs :

- au sentiment de sécurité de l'internaute quant à la transaction bancaire (Bellman *et al.*, 1999; Kehoe *et al.*, 1997; Kehoe *et al.*, 1998a; Kehoe *et al.*, 1998b; Pitkow & Kehoe, 1996),
- à son sentiment quant à la confidentialité des informations personnelles (non divulgation à un tiers de ces dernières ou de celles recueillies automatiquement lors du parcours et des achats effectués sur le site ; enregistrement des fichiers-clients de l'entreprise à la C.N.I.L.¹³⁸ et possibilité pour l'internaute de les modifier) (Teo, 2002; Tilson *et al.*, 1998),

- et à sa satisfaction quant aux facilités de paiement proposées (Nielsen *et al.*, 2001; Teo, 2002).

Les entretiens utilisateurs, divers ouvrages sur la sécurisation sur Internet et les observations comparées de sites marchands (cf. tableau 24 ci- après) nous ont conduit à ajouter des évaluations très particulières concernant le sentiment de sécurité de l'internaute ou à son impression par rapport :

- à l'existence et à la fiabilité de partenaires financiers et d'organismes de contrôle (Fouchard, 2001),
- au non-archivage du numéro de carte bancaire ou le contraire.

Concernant ce dernier point, dans un cas, l'internaute peut passer une commande plus rapidement les fois consécutives et dans l'autre cas, il doit entrer son numéro de carte bancaire à chaque transaction, ce qui est plus contraignant pour lui mais plus sécuritaire aussi. En effet, ce fameux numéro ne figure alors jamais sur les fichiers-clients conservés par l'entreprise. Ces derniers présentent une certaine vulnérabilité puisque plusieurs sites marchands optent pour le non-archivage et qu'un des risques majeurs pour l'entreprise est celui de la pénétration de son environnement informatique à partir de son site Internet (Ghernaouti-Hélie, 2002).

Tableau 24 : Besoins en connaissances des professionnels du Web marchand concernant la Sécurité et le Paiement

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Sentiment de sécurité	Questionnaire+ Outil M.A.S.C. + Littérature/Documentation
Non-archivage du n° CB	Littérature/Documentation
Partenaires financiers & Organismes de contrôle	Entretiens utilisateurs + Littérature/Documentation
Confidentialité des informations personnelles	Questionnaire + Littérature/Documentation
Facilités de paiement	Questionnaire+ Littérature/Documentation + Outil M.A.S.C.

3.2.10. Interactivité

Ayant moins de données des professionnels de la veille-marketing sur ce point, cette catégorie a été inspirée du questionnaire Noël 2001, de plusieurs travaux et surtout de l'outil M.A.S.C. (cf. tableau 25 ci-après). Ce dernier réunit l'interactivité directe, c'est-à-dire le contact (par téléphone ou par messagerie électronique) avec un service clients, aussi bien que l'interactivité en ligne, c'est-à-dire tout ce qui est relatif à la possibilité de communiquer avec d'autres internautes ou d'accéder à des informations supplémentaires depuis le site.

Concernant l'interactivité hors ligne, il faut savoir que le service clients est un critère de différenciation entre les sites de commerce électronique, en raison de la standardisation croissante des produits/services du Web (Macarez & Leslé, 2001). Une fois de plus, les entreprises mettent ainsi en place une structure plus ou moins conséquente pour le traitement des requêtes des clients et il semble nécessaire de vérifier si ces efforts sont suffisamment efficaces aux yeux de leurs clients. D'ailleurs trois questions sont consacrées à ce point dans le questionnaire d'évaluation de Noël 2001.

Dans cette catégorie, nous avons suggéré d'évaluer :

- la proportion d'internautes de la cible clients ayant remarqué que c'est un service proposé par le site et qui ont le réflexe de l'utiliser,
- la satisfaction des internautes concernant les moyens pour contacter le service (adresse courriel, téléphone) (Lee *et al.*, 2000),
- et leur satisfaction concernant la qualité des réponses apportées et éventuellement concernant la convivialité du personnel du service clients (Fouchard, 2001).

D'importants moyens sont également mis en œuvre pour l'interactivité en ligne. Nous avons proposé d'évaluer :

- la satisfaction des internautes quant aux propositions de liens vers des sites d'informations, des chats, des forums (Fouchard, 2001; Lambert & Legrand, 2001; Nielsen *et al.*, 2001; Teo *et al.*, 2003),
- leur satisfaction quant à la possibilité de vote électronique et de lire ou de déposer des témoignages sur les produits du site marchand (Lambert & Legrand, 2001; Lohse & Spiller, 1998b).

Tableau 25 : Besoins en connaissances des professionnels du Web marchand concernant l'Interactivité

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Possibilité de contacter un service clients	Entretiens utilisateurs + Littérature/Documentation
Moyens pour contacter le service	Questionnaire + Littérature/Documentation + Outil M.A.S.C.
Qualité des réponses apportées & Convivialité	Questionnaire + Littérature/Documentation
Service spécial conseil achat	Entretiens prof. VM
Liens vers des sites d'informations, des chats, des forums	Littérature/Documentation + Outil M.A.S.C.
Vote électronique & Témoignages	Littérature/Documentation

3.2.11. Logistique

La première raison invoquée pour acheter par Internet est la commodité, la seconde, le fait de ne pas avoir la pression du personnel de vente et la troisième, le gain de temps (Kehoe *et al.*, 1997; Kehoe *et al.*, 1998a). Les commerces électroniques doivent donc être efficaces et rapides pour répondre à ces attentes d'achat à distance en toute quiétude. Cela concourt à la nécessité de vérifier si les cyber-consommateurs sont satisfaits de la logistique assurée par les commerces électroniques.

Aussi en lien avec les regroupements des professionnels de la veille-marketing et l'outil M.A.S.C. (cf. tableau 26 ci- après), nous avons formé cette catégorie qui propose d'évaluer la satisfaction des cyber-consommateurs concernant :

- l'efficacité de la livraison (respect des délais, arrivée du produit en bon état, etc.),
- les frais d'envoi et la somme minimum d'achat pour une livraison gratuite,
- l'intérêt de l'emballage cadeau,
- et l'efficacité du suivi de commande.

Par ailleurs, nos entretiens utilisateurs, notre propre expérience d'achat par Internet et la prise en compte des différentes caractéristiques logistiques des sites marchands (Macarez & Leslé, 2001) nous ont fait ajouter des évaluations particulières (à ne réaliser que s'il y a une volonté d'amélioration et de transformation en profondeur) pour évaluer la satisfaction des cyber-consommateurs concernant :

- la livraison avec rendez-vous (s'il s'agit d'un transporteur interne/externe pratiquant ce genre de service ou un « e-logisticien » gérant toute la chaîne de distribution depuis le producteur ou fournisseur jusqu'au panier de la ménagère),
- l'intérêt/efficacité des relais colis (si l'entreprise marchande jouit d'un réseau de boutiques assurant ce relais),

- les modalités de retour des produits (emballage intact, facture conservée, renvoi par La Poste généralement à ses propres frais, etc.).

Tableau 26 : Besoins en connaissances des professionnels du Web marchand concernant la Logistique

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Garanties de livraison (Respect délai & Bon état produit)	Entretiens prof. VM + Questionnaire + Outil M.A.S.C. + Littérature/Documentation
Emballage cadeau	Questionnaire
Frais d'envoi & Somme minimum d'achat pour une livraison gratuite	Entretiens prof. VM + Littérature/Documentation + Outil M.A.S.C.
Possibilité de RDV	Littérature/Documentation
Relais colis	Littérature/Documentation
Suivi de commande	Entretiens prof. VM + Littérature/Documentation + Outil M.A.S.C.
Retour de produits	Entretiens utilisateurs + Littérature/Documentation+ Outil M.A.S.C.
Service de maintenance.	Littérature/Documentation

3.2.12. Les informations associées au service de vente

Cette catégorie concerne les pages explicatives sur les conditions de ventes, les garanties de livraison, les coordonnées de divers services, les aides, etc. Les professionnels de la veille-marketing s'intéressent peu à ces informations, peut-être parce qu'ils ne les rédigent pas. Dans le cas du Sitemarchand.com, ce sont les professionnels du marketing ou le responsable du Webmastering qui s'occupent de les formuler. Cependant, dans de nombreux cas, il est conseillé de faire appel à des spécialistes de la rédaction de documents techniques pour le Web (Goto & Cotler, 2002; Nielsen, 2000). Ainsi, c'est l'outil M.A.S.C., quelques travaux et des comparaisons de plusieurs sites marchands qui nous ont conduit à former cette catégorie (cf. tableau 27 ci-après).

Le choix des informations est, en effet, essentiel pour la crédibilité des sites marchands (Jarvenpaa & Todd, 1997; Teo *et al.*, 2003). Un site précise, par exemple, que ses fichiers-clients sont enregistrés à la C.N.I.L. conformément à la loi « Informatique et Liberté » du 6 janvier 1978. Cela n'est pas toujours précisé sur tous les sites. Cette information devient pourtant décisive quand on apprend que seulement 45% des grands sites d'e-business français ont déclaré leurs fichiers (Macarez & Leslé, 2001).

Dans cette catégorie, nous avons proposé l'évaluation de la satisfaction des internautes concernant :

- la qualité et le choix de ces informations,
- leur complétude,
- leur compréhension en termes de vocabulaire,
- et leur visibilité qui renvoie à la catégorie « Look & co ».

Tableau 27 : Besoins en connaissances des professionnels du Web marchand concernant les Informations associées au service de vente

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Qualité & Choix de ces informations	Littérature/Documentation + Outil M.A.S.C.
Complétude de ces informations	Littérature/Documentation
Compréhension	Littérature/Documentation
Visibilité de toutes ces informations	Littérature/Documentation + Outil M.A.S.C.

3.2.13. Facilité à trouver un produit

Les deux professionnels de la veille-marketing trouvent que les trois dimensions « Existence de différentes fonctions de recherche », « Facilité de la recherche » et « Pertinence et facilité de lecture des résultats de la recherche » sont trop nombreuses (dans la liste exhaustive de dimensions utilisateurs). Selon eux, il faut condenser les choses. C'est ce que nous avons fait en réalisant la catégorie « Facilité à trouver un produit », dont le titre est inspiré des propos du premier professionnel de la veille-marketing qui pense que l'internaute préfère trouver un produit plutôt que de le chercher. Certes ! Mais c'est encore loin d'être le cas sur certains sites marchands (Katz & Aspden, 1997; Nielsen *et al.*, 2001; Pitkow & Kehoe, 1996; Spool *et al.*, 1999).

Nous avons suggéré d'évaluer la satisfaction des internautes concernant :

- la facilité générale à trouver les produits, comme dans le questionnaire Noël 2001, puis d'approfondir les éléments suivants, en particulier en cas de mauvais résultat :
 - la facilité d'utilisation et l'efficacité de la recherche par mots-clefs,
 - la pertinence du choix des rubriques et sous-rubriques pour trouver un produit,
 - la facilité à repérer le produit recherché dans les listes de résultats.

Tableau 28 : Besoins en connaissances des professionnels du Web marchand concernant la Facilité à trouver un produit

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Trouver le produit	Entretiens prof. VM + Questionnaire + Littérature/Documentation
Recherche par mots-clefs	Entretiens utilisateurs + Littérature/Documentation + Outil M.A.S.C.
Les rubriques et sous-rubriques	Entretiens utilisateurs + Littérature/Documentation
Listes de résultats	Entretiens prof. VM + Entretiens utilisateurs + Littérature/Documentation

3.2.14. Facilité à commander un produit

En revanche, les deux professionnels de la veille-marketing sont plus préoccupés par la facilité à commander un produit. Concernant le panier, « *du statut de métaphore, il est passé à celui d'éléments d'interface standard* » et « *une fois qu'un élément est suffisamment répandu, il devient une norme, et les utilisateurs savent d'avance ce qu'il faut en attendre et comment il fonctionne* » (Nielsen, 2000, p. 186). Cela est certes juste, cependant, plusieurs études (Nielsen *et al.*, 2001; Pilverdier & Gaillard, 2001; Rehman, 2000) et nos entretiens utilisateurs montrent qu'il existe encore des difficultés relatives à l'utilisation du panier, ainsi qu'au remplissage des bons de commande.

De même que précédemment, nous avons d'abord proposé une question généraliste pour déterminer la proportion d'internautes rencontrant ces problèmes. Cette question figure aussi dans le questionnaire Noël 2001. Si nécessaire, nous avons ensuite proposé d'approfondir l'étude.

Il s'agirait donc d'évaluer la satisfaction des internautes concernant :

- la facilité générale à commander un produit, puis :
 - la facilité de saisie de la taille, de la quantité, etc. des produits (Rehman, 2000),
 - la facilité d'utilisation du panier (Fouchard, 2001; Pilverdier & Gaillard, 2001; Tilson *et al.*, 1998),
 - la facilité de remplissage du bon de commande (Nielsen *et al.*, 2001),
 - et la simplicité du récapitulatif de produits & frais d'envoi (Fouchard, 2001).

Tableau 29 : Besoins en connaissances des professionnels du Web marchand concernant la Facilité à commander un produit

Besoins en évaluation	Sources ayant permis l'identification du besoin en évaluation
Commander le produit	Entretiens prof. VM + Questionnaire+ Littérature/Documentation
Saisie de la taille, de la quantité, etc. des produits	Entretiens utilisateurs + Littérature/Documentation
Panier	Entretiens prof. VM + Littérature/Documentation
Bon de commande	Entretiens prof. VM + Littérature/Documentation
Récapitulatif de produits & frais d'envoi	Entretiens utilisateurs + Littérature/Documentation

4. CONCLUSION

A partir des données des entretiens avec les professionnels de la veille-marketing et du Webmastering, nous avons tout d'abord construit un cahier des charges du futur outil d'aide à la conception d'évaluations. A l'aide de ces données et de celles disponibles dans la littérature ainsi que dans les précédents entretiens utilisateurs et enfin à l'aide de notre propre analyse comparative de plusieurs sites marchands, nous avons identifié les besoins en évaluation des professionnels du Web marchand qui semblent raisonnables. Nous avons alors élaboré une ébauche de la boîte à outils qui détaille ce qu'il s'agit d'évaluer. Elle comporte 13 catégories regroupant plusieurs besoins en évaluation spécifiques. De plus, pour chaque besoin en évaluation, nous avons précisé l'objectif de l'évaluation (« les internautes sont-ils satisfaits... ? »), les techniques d'évaluation adaptées (questionnaire, entretiens individuels ou *focus groups*) et parfois les questions à formuler, voire certains détails (les caractéristiques des sujets à interroger, une liste d'artefacts intervenant généralement dans la perception des internautes, etc.).

Cette ébauche a ensuite été soumise aux jugements de quatre professionnels, dont trois, cette fois-ci, sont directement impliqués dans le marketing et la conception d'un site marchand. En effet, dans le chapitre suivant, poursuivant notre démarche opérationnelle, nous menons des entretiens de validation de l'ébauche de la boîte à outils. Il s'agit de vérifier l'utilité de chacun des besoins en évaluation, la pertinence et la compréhensibilité du contenu et des techniques d'évaluation proposées. Enfin, il s'agit d'examiner si le découpage des 13 catégories correspond à ce que les professionnels s'attendent à y trouver.

Dans le chapitre suivant, nous présentons la boîte à outils dans son ensemble, après y avoir opéré diverses modifications, également explicitées.

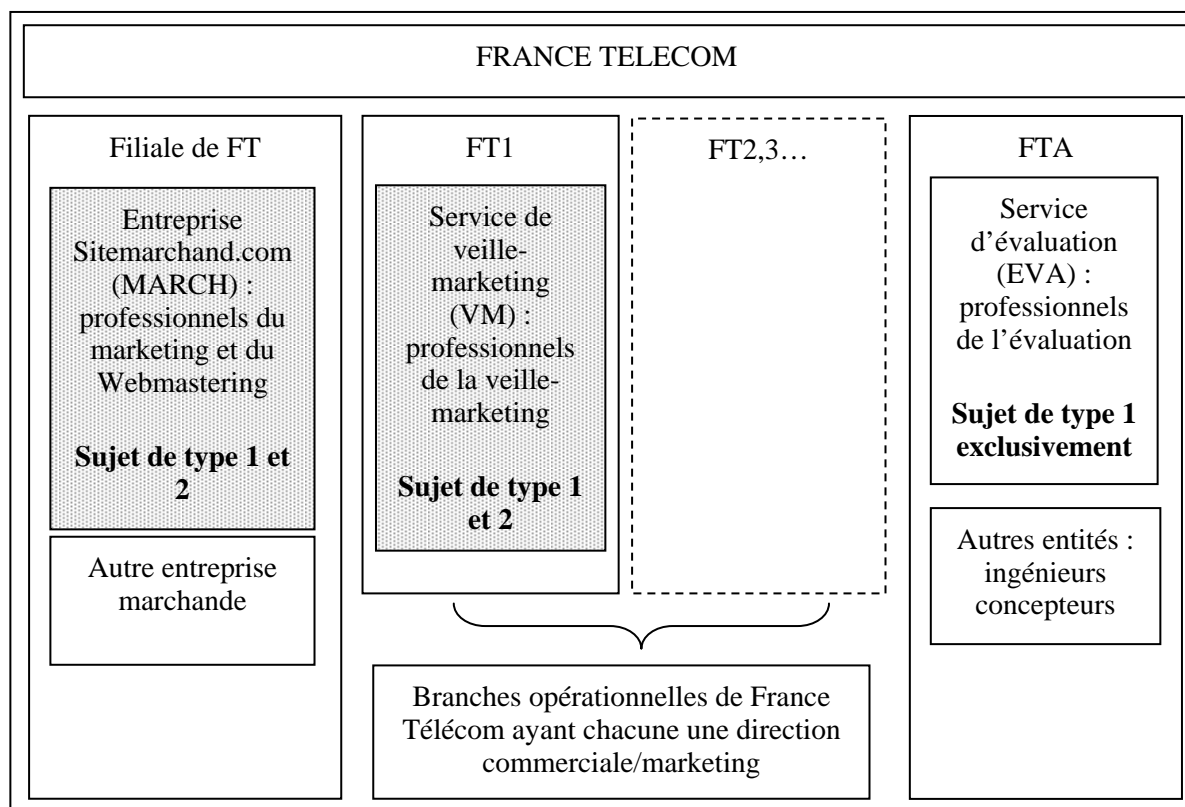
CHAPITRE 6

L'OUTIL D'AIDE A LA CONCEPTION D'EVALUATIONS POUR LES PROFESSIONNELS DU **WEB MARCHAND**

REPERES POUR LE CHAPITRE

Les repères pour situer ce chapitre sont les mêmes que pour le précédent chapitre (cf. figure 17). Nous abordons ici la création de l'outil d'aide à la conception d'évaluations pour les professionnels du Web marchand à partir de l'opérationnalisation des dimensions utilisateurs élaborées au chapitre 3 et insuffisamment transformées au chapitre 4. Les professionnels du Web marchand considérés sont les sujets de type 1 qui élaborent les évaluations et aussi les sujets de type 2 qui transforment le produit/service et ont certains besoins en connaissances, identifiés dans le chapitre 5 qui précède.

Figure 17 : Organigramme simplifié des branches et filiales de France Télécom concernées dans cette thèse



Légende : Les parties grisées correspondent aux services des professionnels auxquels nous nous intéressons dans ce chapitre. Les sujets de type 1 ont une activité d'évaluation et les sujets de type 2 ont une activité de conception/marketing des produits/systèmes.

1. INTRODUCTION

Dans ce chapitre, nous poursuivons la construction de l'outil d'aide à la conception d'évaluations pour les professionnels du Web marchand et décrivons sa genèse, élément par élément.

Nous exposons brièvement le déroulement des entretiens de validation de l'ébauche de l'outil ou boîte à outils, élaborée à partir des données et spécifications du précédent chapitre. Ces entretiens ont été menés avec un des professionnels de la veille-marketing, le responsable du Webmastering et deux autres professionnels de l'entreprise Sitemarchand.com. Nous nous appuyons également sur une autre trace de leur activité, un second questionnaire d'une enquête d'évaluation du Sitemarchand.com, ainsi que sur un ensemble de connaissances¹³⁹ portant sur les activités déployées au sein d'un commerce électronique, d'une part, et sur les usages et la perception des cyber-consommateurs, d'autre part.

Nous détaillons ensuite le contenu de la boîte à outils, élément par élément, ce qui nous permet de justifier chaque besoin en évaluation et de proposer différentes techniques d'évaluation appropriées. Dans ce chapitre, nous développons davantage les réflexions et commentaires des professionnels à partir de leurs verbalisations retranscrites. Les données sur leurs activités et besoins en évaluation sont plus concrètes et plus riches.

Nous rappelons que notre démarche est basée sur un principe opérationnel : chaque élément de la boîte à outils doit être justifié soit par un besoin en évaluation exprimé par les professionnels présentés ci-dessus, soit par les traces de leur activité, soit par des éléments de connaissances issus de la littérature du domaine sur l'activité de tout professionnel du Web marchand. Chaque besoin existant est traduit en une rubrique « Élément à évaluer », détaillant l'« Objectif de l'évaluation » et le « Type et détails de l'évaluation ».

La logique opérationnelle est aussi celle de construire une ébauche et de la modifier si nécessaire. Les entretiens avec les professionnels nous conduisent à accomplir des modifications (a) concernant, non pas le découpage des 13 catégories en lui-même, mais l'ordre de succession des catégories, (b) en termes d'ajout ou de suppression de certains besoins en connaissances spécifiques, (c) en termes de déplacement de certains de ces besoins au sein des 13 catégories et enfin (d) en termes de reformulation de leur intitulé.

Par ailleurs, nos travaux présentent certaines limites. Les professionnels étant très occupés, il n'est pas possible d'obtenir plus d'un entretien avec chacun d'entre eux. Certaines

inférences ont donc été réalisées à partir de leurs verbalisations concernant les exigences de leurs activités. Des entretiens avec des professionnels d'autres services (service logistique, service produits, service clients) auraient aussi permis de consolider nos données. Ainsi, une connaissance plus fouillée des exigences de l'activité de ces divers professionnels conduirait certainement à poursuivre des modifications, de même, a fortiori, s'il était possible de recueillir des données sur l'activité de professionnels d'autres entreprises marchandes. Néanmoins, la boîte à outils présentée ici constitue un aboutissement à partir de la synthèse des données que nous avons pu recueillir sur le terrain et des divers éléments de connaissances.

2. VALIDATION DE L'ÉBAUCHE DE LA BOÎTE A OUTILS

Les entretiens de validation de l'ébauche de la boîte à outils ont été menés avec plusieurs professionnels exclusivement du Web marchand :

- le professionnel de la veille-marketing, spécialiste des sites marchands,
- et trois professionnels de l'entreprise Sitemarchand.com :
 - le précédent responsable du Webmastering,
 - un autre professionnel du Webmastering (le responsable de la charte graphique),
 - et un professionnel du marketing, chargé des études d'évaluations et du marketing de fidélisation des personnes déjà clientes (appelé simplement professionnel du marketing).

Les entretiens semi directifs ont duré de une à trois heures en fonction de la disponibilité des professionnels et se sont déroulés tous de la même façon :

- explication de l'objectif poursuivi avec la construction de la boîte à outils et de l'objectif de l'entretien,
- passage en revue de chaque « besoin en évaluation » susceptible de correspondre à l'activité du professionnel (ex. ceux de la catégorie « Look & co » avec le responsable de la charte graphique) et recueil de leurs commentaires. En revanche, la boîte à outils a été entièrement passée en revue avec le professionnel de la veille-marketing (spécialiste des sites marchands),
- les questions posées et/ou relances utilisées pour recadrer les discussions ou les solliciter sont du type : « Est-ce réellement un résultat en évaluation dont vous vous serviriez ? », « De quoi auriez-vous vraiment besoin comme information sur ce point ? »,

- des discussions sont aussi menées pour éclairer le fonctionnement de l'entreprise Sitemarchand.com puisqu'elles permettent de mieux comprendre l'activité des différents acteurs et donc leurs besoins en connaissances.

Ces entretiens ont été retranscrits et les verbalisations pertinentes des quatre professionnels mobilisées pour modifier ou renforcer chaque élément de la boîte à outils.

De plus, nous avons obtenu un second questionnaire utilisé pour évaluer le Sitemarchand.com, celui de Noël 2002, qui nous a permis d'enrichir la boîte à outils ébauchée. Ce questionnaire nous a été transmis par le professionnel du marketing.

3. PRESENTATION DE LA BOITE A OUTILS VALIDEE

3.1. INTRODUCTION

Nous présentons la boîte à outils après validation, élément par élément, en précisant ce qui a été modifié. Nous avons notamment modifié l'ordre des 13 catégories (cf. tableau 16 du chapitre précédent) pour qu'il corresponde davantage aux représentations des professionnels (au moins globalement, étant donné qu'ils ont des activités différentes). Le nouvel ordre proposé respecte davantage l'importance relative des catégories selon les professionnels interrogés (cf. tableau 30 ci-après). Par exemple, les professionnels du marketing comme du Webmastering se préoccupent relativement peu des aspects sécurité, étant donné que la sécurisation de la transaction est assurée par les professionnels du service informatique. La catégorie « Sécurité & Paiement » a ainsi été placée après celles concernant les aspects logistiques ou d'interactivité entre les internautes et le site marchand.

Tableau 30 : Principaux besoins en connaissances des professionnels du marketing et de la conception pour évaluer les sites Internet marchands

<ul style="list-style-type: none">• Notoriété• Facilité d'accès & Capacité du système• Page d'accueil• Look, présentation, visibilité et navigation • Offre & Positionnement• Description des produits• Netmarketing • Logistique & Services• Interactivité• Sécurité & Paiement • Facilité à trouver un produit• Facilité à commander un produit • Les informations associées au service de vente
--

De même, l'ordre de certains besoins en évaluation au sein de chacune des catégories a été changé et des besoins ont été déplacés vers une catégorie plus appropriée. Nous avons aussi supprimé des éléments irréalistes car ils ne représentaient pas vraiment un besoin en évaluation et en avons ajouté d'autres qui ont émergé lors des entretiens de validation. Nous avons transformé plus ou moins profondément certains besoins en évaluation existants, et notamment rendu plus concret leur contenu (ex. reformulation de questions « pré-proposées » non appropriées). Enfin, nous avons modifié certains intitulés, la plupart du temps sur la base de suggestions explicites des professionnels. Un récapitulatif de tous les besoins en évaluation de la boîte à outils, avant et après validation, présenté dans le nouvel ordre, est détaillé en annexes du chapitre 6.

3.2. PRESENTATION DE LA BOITE A OUTILS ELEMENT PAR ELEMENT

Nous employons les abréviations suivantes dans la boîte à outils :

- Q : questionnaire (qui correspond à une enquête de satisfaction, fidélisation, etc.),
- E : entretiens individuels, avec éventuellement utilisation du site à étudier,
- FG : *focus groups*,
- EA : étude approfondie et/ou spécifique qui concerne certains sites marchands et/ou dans certains contextes et qui supposent généralement une remise en question assez vaste du fonctionnement de l'entreprise marchande ou de ses investissements.

3.2.1. Catégorie « Notoriété »

3.2.1.1. Notoriété

Tableau 31 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Notoriété »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Notoriété	Enquête concurrentielle : Les internautes de la cible clients du site (visiteurs, acheteurs et prospects ¹⁴⁰) connaissent-ils le site par rapport à ceux de la concurrence ?	EA détachée de l'enquête de satisfaction ou de fidélisation sur le site à étudier, car cela instaure un biais : l'internaute sait d'avance qu'il répond à un questionnaire envoyé par le site marchand donné, ce qui peut influencer ses réponses aux questions concernant la notoriété de ce site (et de ses concurrents). Etude par Q ou par E auprès de différents échantillons d'internautes (clients, prospects, etc.) : Appréciation de leurs connaissances du site marchand par rapport à d'autres sites du même secteur de vente. <ul style="list-style-type: none"> • notoriété spontanée : énonciation libre des sites connus, • notoriété assistée : avec présentation d'une liste de sites.

L'évaluation de la connaissance des internautes concernant la notoriété du site au regard des sites de la concurrence est un réel besoin en évaluation. Le professionnel de la veille-marketing le confirme et propose même que ce soit une question incluse dans un questionnaire d'évaluation généraliste. Nous pensons que cela n'est cependant pas réaliste. Si l'évaluation porte sur la notoriété en spontané¹⁴¹ ou en assisté¹⁴², tel que nous le suggère le second professionnel de la veille-marketing lors des premiers entretiens libres (présentés au chapitre 4), alors elle doit se faire en dehors de l'enquête généraliste par questionnaire. En effet, il faut éviter les biais possibles. Or, l'internaute sait qui envoie le questionnaire (a fortiori, si c'est un questionnaire en ligne sur le site). Il s'agit donc d'une évaluation à part, dont la dénomination technique est « enquête concurrentielle ».

3.2.1.2. Image de marque

Tableau 32 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Image de marque »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Image de marque	Quelle est l'image perçue par les internautes visiteurs, acheteurs ou prospects du site étudié ? Quelle(s) composante(s) principale(s) participent à	Etude par Q, détachée ou incluse dans l'enquête généraliste de satisfaction ou de fidélisation. Appréciation (quantitative) de la perception des internautes concernant : <ul style="list-style-type: none"> • l'attractivité du site, • le degré de confiance, • la satisfaction générale quant au site,

	<p>cette bonne ou mauvaise image (ex. « image prix », image qualité des produits, image « innovateur », ...) ?</p> <p>Sur quelle(s) caractéristique(s) (dynamiques, morales, etc.) reposent l'image perçue des internautes ?</p>	<ul style="list-style-type: none"> • puis cette satisfaction est-elle liée plus particulièrement à la qualité des produits, aux prix pratiqués, à la livraison, au S.A.V., etc. (cf. aussi catégories « Offre – Positionnement » et « Logistique »), • diverses caractéristiques perçues – ex. : <ul style="list-style-type: none"> ○ dynamiques (jeune, vieillie), ○ morales (sérieux, qualité), ○ sociales (distinction, banalité), ○ symboliques (quel genre de personne utilise ce site), etc.
--	--	---

L'image de marque constitue aussi un besoin en évaluation. Le professionnel de la veille-marketing nous l'affirme et nous indique encore qu'il est possible d'intégrer les questions s'y rapportant dans un questionnaire d'évaluation plus complet. En effet, cela nous semble possible puisqu'il s'agit alors de questions interrogeant la satisfaction des internautes concernant le site à étudier et non leurs connaissances quant aux sites marchands.

Les différents éléments proposés ici (attractivité, degré de confiance, satisfaction générale, caractéristiques dynamiques, morales, etc.), à évaluer de façon quantitative, sont essentiellement issus de Dayan (2001) et ne sont pas contestés par les professionnels interrogés. Le professionnel du marketing nous précise qu'effectivement ce genre d'études est destiné à confirmer ou à infirmer l'orientation choisie selon « *si on veut une image prix, si on veut une image qualité, une image d'innovateur* ».

3.2.1.3. Efficacité et poids relatif des différents supports médiatiques

Tableau 33 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Efficacité de la médiatisation et poids relatif des différents supports médiatiques »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Efficacité et poids relatif des différents supports médiatiques	<p>Les internautes, nouveaux clients sur le site, ont-ils découvert le site grâce à la médiatisation du site ou autrement ?</p> <p>Combien d'internautes ont découvert le site par chacun des différents supports de médiatisation (presse, radio, etc.) ?</p>	<p>Etude par Q auprès d'internautes nouveaux visiteurs-clients (car les anciens risquent d'avoir oublié comment ils ont découvert le site à l'origine).</p> <p>Proportion d'internautes ayant découvert le site par :</p> <ul style="list-style-type: none"> • la presse, • les annonces radio, • l'affichage, • Internet (cf. aussi la rubrique « Référencement »), • les courriels publicitaires, d'informations, etc. • le bouche à oreille / marketing viral (ex. discussion en ligne), • un bon d'achat. <p>Possibilité aussi de corrélérer :</p> <ul style="list-style-type: none"> • les données mensuelles concernant les quantités de publicités mises en ligne par le site et par ses concurrents, • et le taux d'audience et de transformation du site (cf. « Raisons de retour des acheteurs »). <p>Donc estimation de l'impact des publicités en ligne sur le nombre de visites et de ventes sur le site.</p>

L'évaluation de l'impact de la médiatisation (presse, annonces radio, etc.) sur la découverte du site nous est donnée par le questionnaire d'évaluation de Noël 2001. Les professionnels du (de la veille-) marketing nous le confirment logiquement comme étant un besoin en évaluation. Le professionnel de la veille-marketing conseille cependant d'ajouter « poids relatifs des différents supports médiatiques » à son intitulé, car il ne le trouve pas assez explicite.

L'évaluation par questionnaire proposée peut aussi être corrélée à des données objectives. Le professionnel du marketing évoque, en effet, la possibilité de comparer les quantités de publicités mises en ligne mensuellement par le Sitemarchand.com et ses sites concurrents et de voir si le taux d'audience augmente simultanément (« *Si on veut aller plus loin il faut arriver à reconnaître... Si un mois on est plus reconnu que les autres, si c'est parce que nous on a plus communiqué que les autres.* »). Il peut être observé, de même, si le taux de transformation¹⁴³ augmente, puisqu'il s'agit d'une métrique importante qui permet de mieux connaître la performance d'un site marchand (Kim & Lee, 2002; Rehman, 2000).

Cependant, le professionnel du marketing n'est pas spécialiste de la communication et de la constitution du « média-planning ». Il serait donc utile d'approfondir ce point et le suivant (cf. « Réception des messages publicitaires ») avec de tels professionnels, mais cela n'a pas été possible. Notamment, nous aurions voulu en apprendre davantage sur leurs outils et méthodes d'évaluation des actions de la médiatisation « offline » (Dayan, 2001). Nous avons ainsi abrégé les aspects « médiatisation » en supprimant l'élément « Efficacité de la médiatisation « offline » du site sur ses ventes et sa notoriété » (cf. annexes chapitre 6).

3.2.1.4. Raisons de retour des acheteurs sur le site

Tableau 34 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Raisons de retour des visiteurs sur le site »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Raisons de retour des acheteurs sur le site	<p>Enquête fidélisation : Pourquoi les internautes, déjà clients du site, sont-ils retournés sur le site ?</p> <p>Quel est l'impact de l'image de marque (ou en tout cas de la satisfaction des clients) et de diverses opérations marketing ?</p>	<p>Etude par Q (= enquête de fidélisation). Evaluation de la proportion d'internautes étant revenus pour les raisons suivantes :</p> <ul style="list-style-type: none"> • par satisfaction des précédents achats (possibilité d'approfondir ce point, cf. « Image de marque », ci-dessus), • une publicité en ligne concernant le site (ex. produit vendu par le site), • un courriel avec des offres, • un bon d'achat (envoyé par courriel ou autre), • par habitude. <p>Corrélation avec les mesures des flux d'audience recueillies à partir d'un panel et à partir d'un outil de « tracking » interne au site.</p>

Cette enquête de fidélisation (pourquoi les clients sont-ils retournés sur le site ?) et son contenu (par satisfaction des précédents achats, une publicité en ligne, etc.) sont aussi inspirés du questionnaire de Noël 2001. Le professionnel de la veille-marketing nous atteste qu'il s'agit d'un besoin en évaluation complémentaire au précédent. Une fois de plus, il en réfute le titre. Pour lui, il s'agit seulement des raisons de retour des internautes sur le site (nouveau titre) et l'ancien titre « Médiatisation / Image de marque et raisons de retour des visiteurs sur le site » n'est pas approprié (cf. annexes du chapitre 6).

L'enquête de fidélisation proposée peut aussi être corrélée à des données objectives. Le précédent professionnel nous précise que les mesures des flux d'audience recueillies à partir d'un panel¹⁴⁴ et/ou d'un outil de « tracking »¹⁴⁵ au niveau du serveur du Sitemarchand.com (cf. aussi Macarez & Leslé, 2001) permettent de savoir à une plus grande échelle par quel lien

les internautes sont arrivés sur le site (un moteur de recherche, lequel ?, un lien d'affiliation, une publicité en ligne, etc.).

3.2.1.5. Réception des messages publicitaires

Tableau 35 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Réception des messages publicitaires »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Réception des messages publicitaires (« offline »)	Concernant les différents messages publicitaires (annonces radio, affiches, publicités en ligne, etc.), les internautes (prospects ou clients) ont-ils retenu l'émetteur du produit/service promu ? Les internautes ont-ils retenu qu'il s'agissait du site étudié ou d'un site marchand concurrent ?	EA. Estimation de la proportion d'internautes ayant : <ul style="list-style-type: none"> • mémorisé l'émetteur (la marque ou le nom du site marchand) versus le type de produit/service, • mémorisé l'émetteur versus un autre émetteur (une marque au lieu d'une autre ou un nom de site au lieu d'un autre). Cela concerne les professionnels réalisant le plan de médiatisation ou le média-planning

Nous avons conservé, plus à titre indicatif, cet élément inspiré par Dayan (2001), qui porte sur la réception des messages publicitaires. Les détails de ce genre d'évaluations seraient à approfondir à partir d'une meilleure connaissance de l'activité et des outils des professionnels spécialistes de la communication et de l'élaboration du « média-planning ».

3.2.2. Catégorie « Facilité d'accès & Capacité système »

3.2.2.1. Référencement et liens d'accès au site

Tableau 36 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Référencement et liens d'accès au site »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Référencement et liens d'accès au site	Les internautes, nouveaux clients d'un site, ont-ils découvert le site grâce à son référencement sur les outils de recherche Internet ? Eventuellement lequel ?	<p>Etude complémentaire à la médiatisation (par Internet). Evaluation grâce aux données de mesure des flux d'audience de la proportion d'internautes ayant découvert le site par :</p> <ul style="list-style-type: none"> • un lien d'affiliation depuis un autre site, • des publicités en ligne (lesquelles ?), • des outils de recherche Internet (lequel ? : moteur de recherche ex. Altavista, Hotbot ou annuaires ex. Yahoo !, Nomade). <p>Ces résultats concernent les professionnels :</p> <ul style="list-style-type: none"> • du média-planning (modification des publicités en ligne), • du service partenariat (programmes d'affiliation entre sites marchands, révision des liens défectueux ou pas assez visibles), • et du référencement (optimisation du référencement par mots-clefs, inscription sur les annuaires en ligne et vérification du positionnement du site sur les différents outils Internet). <p>Analyse corrélative de l'évolution du chiffre d'affaires pour réajuster les actions médiatiques, de référencement, de partenariat, etc.</p>

D'après les deux professionnels du (de la veille-) marketing, il semble plus logique de travailler avec les mesures des flux d'audience pour connaître la proportion d'internautes ayant découvert le site par un lien d'affiliation depuis un autre site, par des publicités en ligne ou par des outils de recherche.

Le professionnel de la veille-marketing nous précise ensuite quels sont les professionnels et services concernés, si les résultats se révèlent décevants. Les professionnels qui réalisent le « média-planning » (ex. service communication) seront en charge de revoir le « média-planning », donc le choix des publicités du Sitemarchand.com placées sur d'autres sites (partenaires), voire leur design graphique. Les professionnels du service partenariat réviseront leurs accords avec les sites partenaires et surtout, après analyse des liens d'affiliation à partir de ces sites, préciseront les efforts à réaliser : liens défectueux à mettre à jour, liens à mettre en valeur visuellement, etc.

De plus, l'activité de référencement est également à revoir si un résultat s'avère mauvais. Un poste à temps plein est consacré à l'achat aux enchères de mots-clefs relatifs au contenu du Sitemarchand.com, à l'inscription du site sur les annuaires et enfin à vérifier le placement du site sur ces outils de recherche. En effet, cette activité de référencement est cruciale : 20% des internautes seulement consultent les résultats au-delà de 2 pages (Macarez & Leslé, 2001).

Ces résultats doivent être complétés par une analyse de l'évolution du chiffre d'affaires et des dépenses pour ajuster les actions médiatiques, de référencement ou de partenariat (Fouchard, 2001).

3.2.2.2. Rapidité de chargement

Tableau 37 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Rapidité de chargement »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Rapidité de chargement	Les internautes de la cible clients du site sont-ils satisfaits de la disponibilité et de la rapidité de chargement du site (obtention de la page d'accueil) puis des autres pages ?	<p>Connaître l'équipement des internautes (types de connexions, navigateurs, plateformes).</p> <p>Etude plutôt par Q. Appréciation générale de la disponibilité & rapidité d'accès au site (très satisfait, satisfait, etc.).</p> <p>EA avec des internautes experts, appréciation:</p> <ul style="list-style-type: none"> • de la rapidité de chargement (rapide, lent, très lent), d'une part, • et d'autre part, de la disponibilité du site dans la durée (le site vous semble-t-il toujours accessible de la même façon ? pas toujours, souvent, peu souvent, etc. ?). <p>Ou mesure d'indicateurs objectifs de durée de connexion grâce à un scénario de simulation de connexion au site marchand à partir de différentes plateformes (PC, Macintosh), rejoué pendant plusieurs semaines.</p> <p>Indicateurs qui jouent aussi dans la perception des internautes concernant ces temps d'attente (éléments supplémentaires à prendre en compte dans une évaluation plus sophistiquée ou à étudier en tant que pistes d'amélioration) :</p> <ul style="list-style-type: none"> • leur tolérance est plus grande s'il s'agit du chargement d'une page dynamique (nécessitant l'accès à la base de produits, le calcul du panier, etc.) que s'il s'agit d'une page statique (surtout si elle a déjà été chargée auparavant), • leur tolérance est plus grande à certains moments de la journée (« heures de pointe ») ou si le site est perçu comme ayant un haut niveau de trafic, mais pas s'il est perçu comme prospère (car il a alors les moyens d'une plus grande capacité technique), • la barre de progression du chargement et/ou le chargement progressif des pages diminuent les frustrations liées à l'attente.

Nous proposons une évaluation par questionnaire, généraliste, interrogeant la satisfaction globale des internautes concernant la disponibilité et rapidité d'accès au site, comme dans le questionnaire de Noël 2001. Cette évaluation est validée.

Par contre, concernant l'évaluation plus approfondie sur la rapidité de chargement (rapide, lent, très lent), d'une part, et la disponibilité du site dans la durée (le site vous semble-t-il toujours accessible de la même façon ?), d'autre part, le professionnel de la veille-marketing nous fait remarquer, à juste titre, que ce genre d'évaluation doit être limité à des internautes experts, capables de faire cette distinction dans leur activité d'utilisation. De toute façon, dans la plupart des cas, une enquête à grande échelle avec la seule question concernant la « disponibilité & rapidité » sera suffisante.

Il faut ajouter que des indicateurs de temps nécessaire pour établir une connexion peuvent être mesurés en continu, grâce à un scénario de simulation de connexion au site marchand à partir de différentes plateformes (PC, Macintosh), rejoué pendant plusieurs semaines. Cela permet d'obtenir des données quantitatives objectives qui répondent à la question de la disponibilité du site dans la durée. Ces mesures ne sont pas évoquées par les professionnels interrogés, mais sont réalisées par le service d'évaluation. Nous avons ajouté cette indication.

Enfin, si les résultats de cette évaluation s'avèrent mauvais, il faut savoir que les internautes n'ont pas les mêmes jugements et réactions en termes de tolérance, selon la représentation qu'ils ont de la page Internet à télécharger (ex. si la tâche requiert d'accéder à une base de données, si elle implique de faire un calcul ou, au contraire, si elle requiert d'accéder simplement à une page statique déjà téléchargée auparavant). Leur tolérance est cependant toujours améliorée si une barre progressive leur indique l'évolution du chargement ou si le chargement des pages est progressif (Bouch *et al.*, 2000; Dellaert & Kahn, 1999). Ces éléments ont été ajoutés dans la boîte à outils comme des indicateurs de la perception des internautes. Il s'agit soit d'éléments supplémentaires à prendre en compte dans une évaluation plus sophistiquée, soit de pistes d'amélioration sur lesquelles les professionnels peuvent travailler.

3.2.2.3. Simplicité de l'adresse URL

Tableau 38 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Simplicité de l'adresse URL »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Simpli- cité Adresse URL	Les internautes de la cible clients du site mémorisent-ils l'adresse facilement ? Quelles erreurs réalisent-ils ?	EA a priori, seulement bien avant le lancement d'un nouveau site marchand. Etude par Q ou par E individuels (ex. avec une phase d'écriture de l'adresse par le client lui-même). Estimation : <ul style="list-style-type: none"> • Nombre d'internautes utilisant l'adresse du site directement/ nombre d'internautes utilisant les outils de recherche Internet, • Parmi le premier groupe : nombre d'internautes éprouvant des difficultés à se souvenir de l'adresse, • Types d'erreurs réalisées et nombre d'internautes les réalisant.

Les professionnels du (de la veille-) marketing nous le signalent tous les deux : il s'agirait d'une évaluation seulement dans le cas du lancement d'un nouveau site qui veut vérifier si son adresse ne pose pas de problèmes (après, les internautes se sont déjà familiarisés avec l'adresse). Pour porter ses fruits, cette évaluation doit même être réalisée avant d'avoir décidé du nom et/ou de l'adresse URL. Le choix de ces éléments est, en effet, contraint par sa disponibilité sur le marché des noms de marques et de noms de cyber-domaines (Fouchard, 2001; Macarez & Leslé, 2001). Ainsi, une fois le nom et/ou l'adresse URL de l'entreprise enregistrés, il existe très peu de marges de manœuvre, même s'il s'avère que l'adresse URL du site est finalement difficile à mémoriser et à orthographier pour les internautes.

Il ne s'agit d'un besoin en évaluation que pour des nouveaux « pure players » ou pour les entreprises « terrestres » ayant des projets de développement de site Internet.

3.2.3. Catégorie « Page d'accueil »

3.2.3.1. Compréhension immédiate de l'objectif du site

Tableau 39 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Compréhension immédiate de l'objectif du site »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Compréhension immédiate de l'objectif du site	<p>Les internautes comprennent-ils, en un clin d'œil, ce qu'ils peuvent faire ou acheter sur le site ?</p> <p>Savent-ils, dès la page d'accueil, si le site correspond à leurs attentes ou si le site s'adresse à une autre cible clients ?</p>	<p>Etude par Q ou plus approfondie par FG ou E avec utilisation du site, auprès de nouveaux visiteurs ou clients d'un site marchand (ou anciens visiteurs ou clients, s'ils se souviennent de leur première impression).</p> <ul style="list-style-type: none"> • Appréciation de leur compréhension immédiate de ce qu'ils peuvent faire sur le site (question oui/non), • Et/ou recueil de leurs diverses impressions générales (questionnement ouvert), • Appréciation de leur impression quant à l'adéquation du site par rapport à leurs besoins (est-ce que vous avez, tout de suite, eu le sentiment que le site vous concernait ?) <p>EA avec recueil de leur représentation quant à la cible à qui s'adresse la page d'accueil (ex. E individuel : identification des éléments sur la page d'accueil qui leur permettent de déduire à qui s'adresse le site ; ou FG avec test projectif : selon eux, la gamme de produits, les marques, le niveau des prix, l'ambiance générale de la page d'accueil correspondent à quel type de consommateur ?)</p> <p>Liste à titre indicatif :</p> <ul style="list-style-type: none"> • fonction de recherche par prix (sur certains sites – ex. de vin – cette fonction est perçue comme s'adressant à des personnes ayant des petits revenus), • présence de liens (ex. sites de vin) qui proposent une recherche de produits pour les non-initiés ou au contraire pour les connaisseurs, etc.

Le professionnel de la veille-marketing réaffirme ce qu'il a déjà soutenu lors des précédents entretiens : c'est un besoin en évaluation.

Les trois dimensions à évaluer (compréhension immédiate, impressions générales, impression quant à l'adéquation du site par rapport à leurs besoins) dans une enquête par questionnaire semblent lui convenir également. Concernant ce point, il soutient : « *Ca [impression quant à l'adéquation du site par rapport à leurs besoins] ça demande déjà de rester un certain temps sur la page d'accueil, c'est comme les diverses impressions... Tu l'as mis dans le bon ordre. Je trouve que diverses impressions ça correspond effectivement : on a un sentiment que c'est facile, que c'est compliqué, que c'est fouillis, que c'est beau, tout ça... alors que la représentation, en effet, ça demande une reformulation dans la tête. Ca, c'est plutôt pour les gens comme ceci ou comme cela, c'est plutôt une progression ...* ».

Les possibilités d'évaluations plus approfondies par entretiens individuels ou par *focus groups* sont aussi validées. En effet, suite à l'évaluation par questionnaire, nous avons proposé une évaluation approfondie, par entretiens avec mise en situation d'utilisation ou par *focus groups* avec tests projectifs, pour étudier leurs représentations concernant la cible à qui s'adresse le site et les éléments de l'interface qui ont participé à la formation de leurs représentations.

Enfin, les éléments de la « liste à titre indicatif » (cf. tableau ci-dessus) ont été identifiés lors des entretiens utilisateurs, surtout lors des *focus groups* où les participants faisaient des commentaires sur les liens ou fonctions proposés dès la page d'accueil leur donnant une indication sur : à qui s'adresse le site ?

3.2.3.2. Réassurance

Tableau 40 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Réassurance »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Réassurance	<p>Les éléments présents sur la page d'accueil participent-ils à rassurer les internautes ?</p> <p>Les éléments présents sur la page d'accueil participent-ils à augmenter le sentiment de fiabilité de l'information et du contenu du site ?</p> <p>Quels éléments précisément ont participé à cela ?</p>	<p>Etude par Q ou par E avec utilisation du site :</p> <ul style="list-style-type: none"> • Appréciation générale, après avoir observé la page d'accueil, du sentiment : <ul style="list-style-type: none"> ○ de sécurité, ○ quant à la crédibilité de l'information, en particulier sur le contenu en termes d'offres (y compris impression de sérieux liée à la qualité graphique de la page d'accueil), • Questionnement explicite de quel élément présent sur la page d'accueil participe le plus à sécuriser l'internaute. <p>Si évaluation par E avec utilisation du site : recueil des réactions et impressions face aux divers éléments de la page d'accueil remarqués par l'internaute. (Si pbl de visibilité des éléments, cf. catégorie « Look & co »).</p> <p>Liste à titre indicatif d'éléments de réassurance :</p> <ul style="list-style-type: none"> • coordonnées, références de l'entreprise (ex. groupe commercial, n° SIRET, chiffre d'affaire, adresse, téléphone), • logos des partenaires financiers & organismes de contrôle, • icône de cadenas avec message de sécurité, • date de dernière mise à jour du site, • explications succinctes ou slogans particuliers, • graphisme « professionnel » (pas un site pirate).

Les deux professionnels du (et de la veille-) marketing confirment ce besoin en évaluation, même si l'ensemble des professionnels du Sitemarchand.com est surtout préoccupé du sentiment de sécurité de l'internaute à partir du moment où il entre dans le processus de commande (à partir du panier). La décision de ne pas acheter peut pourtant être

liée à un grand nombre de facteurs (Helander & Khalid, 2000). De plus, le sentiment de sécurité est un sentiment latent qui se construit chez l'internaute bien avant de venir sur le panier et même avant de venir sur le site (importance de la réputation du commerçant pour accepter de transmettre ses coordonnées bancaires ; Kehoe *et al.*, 1998a), même si l'internaute se pose, bien sûr, plus concrètement la question de la sécurité au moment du processus de commande.

Nous avons proposé de jauger deux éléments conjointement : le sentiment général de sécurité à partir de la page d'accueil, puis l'impression de sérieux ou de crédibilité que donne la page d'accueil.

Nos entretiens utilisateurs ainsi que d'autres travaux (Fouchard, 2001; Teo, 2002) nous conduisent à penser que tout le site et la page d'accueil en particulier doivent faire en sorte que l'internaute se sente en confiance. Divers éléments présents sur la page d'accueil peuvent contribuer à cela : coordonnées, références de l'entreprise, logo des partenaires financiers, etc. (cf. « liste à titre indicatif » dans le tableau ci-dessus, extraite de nos observations empiriques). Le sentiment de sécurité est aussi lié à l'impression de sérieux que donne l'aspect graphique et à l'impression de crédibilité que donne l'information affichée, en particulier sur le contenu en termes d'offres (Nielsen, 2000; Nielsen *et al.*, 2001; Spool *et al.*, 1999 et nos observations empiriques).

Enfin, nous avons aussi proposé de chercher à savoir quels éléments ont engendré le sentiment de sécurité et de sérieux ou le contraire. Le professionnel du marketing certifie qu'il s'agit d'un besoin en évaluation : « *Ben oui, si le client était rassuré sur le site et qu'est-ce qui faisait qu'il l'était, oui tout à fait* ». Une évaluation par entretiens individuels avec utilisation du site marchand a été suggérée, car elle peut s'avérer très pertinente dans ce cas.

Une autre rubrique traite la question du sentiment de sécurité de l'internaute sur le site et au moment de la transaction en particulier dans la catégorie « Sécurité - Paiement » (voir plus loin). Cette diversité offre aux professionnels la possibilité de choisir d'évaluer l'un ou l'autre aspect (ou les deux), selon leur besoin. Par exemple, l'évaluation ci-dessus est particulièrement indiquée suite à une modification de la page d'accueil.

3.2.3.3. Rubriques & Services principaux

Tableau 41 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Rubriques & Services principaux »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Rubriques & services principaux	Les internautes ont-ils l'impression d'avoir facilement accès aux principaux services et rubriques, dès la page d'accueil ?	<p>Etude par Q auprès d'internautes (a priori non novices). Appréciation de leur connaissance quant au fait d'avoir accès à toutes les rubriques de produits et aux services principaux, depuis la page d'accueil (Avez-vous remarqué avoir accès au panier, au suivi de commande, à votre sélection de produits en attente, à l'aide en ligne, etc. depuis la page d'accueil ? ex. Un libellé « Espace clients » ne permet pas forcément de comprendre qu'on aura alors accès au suivi de commande).</p> <p>Cf. aussi catégories « Netmarketing – Compréhension de la page "compte perso" », « Logistique & Services – Utilité des services », « Facilité à trouver un produit – Rubriques et sous-rubriques».</p> <p>Rq : évaluation d'autant plus indiquée que la page d'accueil a été modifiée.</p> <p>EA par E après utilisation-exploration de la page d'accueil :</p> <ul style="list-style-type: none"> • vérification de l'adéquation, pour les internautes, des libellés des liens d'accès aux rubriques ou aux services avec ce, à quoi ils permettent effectivement d'accéder, • compréhension des diverses difficultés rencontrées pour accomplir certaines actions (ex. accéder au suivi de commande, à la sélection de produits en attente, etc.), • si pbl de visibilité/navigation cf. catégorie « Look & co ».

Le professionnel de la veille-marketing valide cet élément. L'idée est ici d'évaluer si les internautes savent, s'ils ont remarqué qu'ils peuvent accéder à la plupart des services (suivi de commande, panier, aide en ligne, etc.) ou aux rubriques de produits depuis la page d'accueil. La page d'accueil doit, en effet, être un point de repère, que ce soit le point de départ des internautes ou que, perdus, ils y soient retournés par dépit (Nielsen, 2000). Or, nos entretiens utilisateurs ont révélé certaines difficultés qu'ont les internautes à accéder aux différents services (ex. faire une recherche à partir du numéro de catalogue papier, retrouver leur suivi de commande, une présélection de produits en attente, etc.).

Cette évaluation va logiquement de paire avec celle de la compréhension des internautes quant aux libellés des liens d'accès aux rubriques ou services et de leur adéquation, avec ce à quoi ils permettent effectivement d'accéder (Nielsen, 2000). Autrement dit, à partir des libellés donnés sur la page d'accueil, les internautes comprennent-ils qu'ils peuvent rechercher un produit à partir de son numéro de catalogue papier, modifier leurs données

personnelles, accéder à leur suivi de commande, accéder à une aide en ligne, etc. Ces éléments de compréhension des libellés sont abordés par le responsable de la charte graphique, quand il évoque les « pages perso » du Sitemarchand.com¹⁴⁶ (cf. « Netmarketing – Compréhension de la page «compte perso» »), et nous les avons généralisés à la page d'accueil.

Pour évaluer la compréhension des libellés, nous avons proposé une évaluation approfondie, où le recours à des entretiens avec mise en situation d'utilisation est particulièrement indiqué. Ces derniers permettent aussi de comprendre quelles sont les éventuelles lacunes graphiques (ex. boutons pas assez visibles) et les diverses difficultés rencontrées par les internautes (les évaluations approfondies proposées dans les catégories « Look, présentation, visibilité, navigation générale », « Facilité à trouver un produit » complètent celle-ci).

3.2.4. Catégorie « Look, présentation, visibilité et navigation »

Jusque là, l'ordre des catégories n'a pas changé. En revanche, la catégorie « Look, présentation, visibilité et navigation » a été placée avant toutes les suivantes, car elle peut se « décliner » sur la plupart des éléments compris dans les catégories suivantes. Il s'agit en quelque sorte d'une méta catégorie, étudiant les aspects visuels et graphiques, et ce qu'ils engendrent au niveau de la navigation, concernant aussi bien les écrans de l'interface que les courriels envoyés aux clients.

Ce nouveau découpage permet aussi de rassembler les catégories « Offre commerciale », « Description des produits » et « Netmarketing » qui vont « bien ensemble », car elles considèrent toutes des dimensions plus commerciales.

3.2.4.1. Reconnaissance visuelle

Tableau 42 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Reconnaissance visuelle »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Reconnaissance visuelle	Les internautes reconnaissent-ils facilement le site/l'entreprise, la marque ?	<p>Appréciation de la facilité à reconnaître l'entreprise ou la marque, dès la première visite (par Q ou par E avec utilisation du site).</p> <p>Concerne particulièrement : la page d'accueil, les pages d'accueil de chaque famille de produits, les outils de la recherche de produits, le panier, les « pages perso » (création de compte, services personnels).</p> <p>EA en cas de mauvais résultat (E avec utilisation du site) :</p> <ul style="list-style-type: none"> • pour déterminer si le problème est relatif : <ul style="list-style-type: none"> ○ à la reconnaissance du logo (couleurs, taille, trop grande richesse graphique diminuant sa lisibilité, etc.), ○ à l'ambiance graphique générale du site (ex. les couleurs, les animations/dessins, les graphismes, etc. ne correspondent pas assez aux diverses formes de communication sur la marque, l'entreprise marchande), • pour comprendre précisément quelles impressions produisent le logo ou l'apparence visuelle du site. <p>Extension à la reconnaissance sonore si le site possède du son (ex. jingle).</p>

Nous avons mis la « Reconnaissance visuelle » en premier dans la catégorie « Look & co » étant donné qu'il s'agit d'un besoin en évaluation primordial pour le responsable de la charte graphique. Ses propos sont, en effet, éloquents : « *Oui c'est bien ça. Ca, par contre. Ca, c'est à développer même. Je sais pas comment, mais c'est une partie très intéressante [...]. Ca c'est la question de tout ce que vous m'avez fait, c'est la question la plus importante... [...] Moi je prends un autre site [site A], par exemple, qui est horrible. C'est un truc horrible. [...]. Mais ils ont une identité visuelle très claire et dès qu'on tombe sur [le site A], on sait ce que c'est. On sait que c'est eux.... Et, en fait, ce que nous on recherche à [Sitemarchand.com] c'est d'avoir ça. [...] Moi, c'est mon objectif numéro un. C'est, même si vous n'aimez pas le orange parce qu'il y aura toujours des gens qui n'aimeront pas le orange [...], même si vous n'aimez pas le orange, moi ce que je veux c'est que vous sachiez que vous êtes sur [Sitemarchand.com] ».*

Ainsi, nous avons proposé une évaluation générale de la facilité à reconnaître l'entreprise ou la marque dès la première visite, puis une évaluation approfondie, en cas de mauvais résultat, pour comprendre de quelle nature est le problème ;

- pour comprendre ce que les internautes ne reconnaissent pas :
 - l'ambiance graphique du site ne correspond-il pas suffisamment aux diverses formes de communication sur la marque, sur l'entreprise marchande, sur les magasins réels ?
 - la taille, les couleurs, la trop grande richesse graphique du logo diminuent-ils sa lisibilité ?
- pour comprendre précisément quelles impressions produisent le logo ou l'apparence visuelle du site sur les internautes.

En effet, le logo doit rappeler la marque « terrestre » existante (s'il y en a une) et véhiculer l'image de la cyber-entreprise (Fouchard, 2001).

Nous avons proposé des entretiens individuels avec utilisation du produit qui sont indiqués pour ce genre d'évaluations approfondies. Enfin, le responsable de la charte graphique nous précise quelles pages sont particulièrement concernées par cette problématique de la reconnaissance visuelle : la page d'accueil, les pages d'accueil de chaque famille de produits, les outils de la recherche de produits, le panier, les « pages perso ».

3.2.4.2. Attractivité & Esthétique

Tableau 43 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Attractivité & Esthétique »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Attractivité & Esthétique	<p>Les internautes trouvent-ils le site attractif?</p> <p>Les internautes trouvent-ils le site ludique/sérieux, contemporain/ringard, beau/pas beau, etc. ?</p>	<p>Appréciation générale (par Q ou par E avec utilisation du site) :</p> <ul style="list-style-type: none"> • de l'attractivité du site quant à son aspect graphique (très attractif, attractif, peu, etc.), • de diverses dimensions (ludique/sérieux, contemporain/ringard, etc.) pour situer l'ambiance que dégage le site, • de l'esthétique du site (beau/pas beau). <p>Extension à l'ambiance sonore si le site possède du son.</p> <p>Rq : il est inutile d'évaluer si les internautes apprécient le choix précis des couleurs car ces réponses, liées aux goûts personnels de chacun, peuvent être toutes différentes et les concepteurs ne peuvent rien en faire.</p>

L'élément « Esthétique » a été considérablement réduit et rassemblé avec l'élément « Attractivité ». En effet, le responsable de la charte graphique explicite ceci : « J'ai

beaucoup de valideurs, j'ai beaucoup de gens qui ont des positions influentes sur le look de [Sitemarchand.com] et quand je vois déjà les disparités entre ces individus là, parce que ça va être : moi je n'aime pas cette couleur, moi j'aime bien cette couleur, avec un espèce de choix très subjectif, plus de l'ordre du goût. J'aime ou je n'aime pas. Moi c'est des informations qui globalement ne peuvent pas me faire avancer au niveau de l'identité visuelle. Maintenant, c'est toujours ... là où c'est intéressant.... s'il y a, je ne sais pas, je dis, 90% des gens qui disent que c'est pas beau, oui, je vais en tirer des informations ou au moins des directives...mais maintenant cibler sur les couleurs exactes, je sais pas si...je ne pourrais pas certifier que ce soit des infos qui puissent me... ». Il est donc clair que les professionnels du Webmastering ont besoin de savoir seulement si le site est perçu comme globalement beau ou le contraire, mais ils n'ont pas besoin que les internautes soient interrogés, plus en détail, sur leurs préférences en termes de couleurs et d'autres éléments graphiques précis. Ils ne pourront rien faire de ces résultats. Ainsi, une seule dimension a été conservée pour recueillir une indication générale concernant l'esthétique (beau – pas beau).

En revanche, évaluer l'attractivité exercée par le site Internet sur les internautes est quelque chose d'intéressant pour les professionnels du Webmastering : « *Ca [attractif - pas attractif] je pense que c'est un peu plus intéressant parce que peut-être, ça nous permettrait de voir quel niveau ... comment on est considéré par rapport à une identité [visuelle] qui existe ou qui n'existe pas [sur le Sitemarchand.com]* ». En effet, lors de la conception d'une nouvelle version d'un site marchand, il s'agit de poser les questions suivantes : « *How appealing is the new design ? Does it communicate the intended character ?* » (Hassenzahl & Trautmann, 2001, p. 167).

Pour cette évaluation, nous avons proposé plusieurs dimensions : attractif - pas attractif, mais aussi ludique - sérieux, contemporain - ringard, etc. Ces dernières ont été identifiées lors de nos précédents entretiens utilisateurs et se retrouvent dans certains travaux (Kim & Moon, 1998) ou peuvent être complétées par d'autres (Park *et al.*, 2004).

3.2.4.3. Graphismes & Animations

Tableau 44 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Graphismes & Animations »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Graphisme & Animation	<p>Les graphismes, photos, animations ou illustrations diverses sont-elles de bonne qualité, d'après les internautes ?</p> <p>Les internautes sont-ils gênés par les animations du site ?</p>	<p>Appréciation (par Q ou par E avec utilisation du site) :</p> <ul style="list-style-type: none"> de la qualité des graphismes (cela fait référence de façon globale à la finesse des dessins, à la qualité graphique des photos, etc.), de l'éventuelle gêne occasionnée par les animations. <p>Liste non-exhaustive :</p> <ul style="list-style-type: none"> logos, dessins/animations, photos des produits illustrations diverses (cartes, 3D).

Le responsable de la charte graphique valide cet élément sans ajouter aucun commentaire. Nous en avons déduit qu'il s'agit d'un besoin en évaluation, mais qui n'est pas principal. Nous avons proposé de vérifier, de façon globale, l'effet du graphisme : de ce point de vue, le site apparaît-il de qualité ou non aux internautes ? En effet, les sites engendrent différentes impressions subjectives selon leur densité graphique, et ceux ayant une apparence graphique très professionnelle sont les plus appréciés (Spool *et al.*, 1999).

De nombreux commentaires de nos entretiens utilisateurs, corroborés par les travaux de Spool *et al.* (1999), nous ont fait ajouter une évaluation concernant la gêne que peut occasionner les animations (cf. aussi la rubrique « Publicité en ligne » plus loin).

3.2.4.4. Aération

Tableau 45 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Aération »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Aération	Les internautes trouvent-ils les informations suffisamment aérées ?	Appréciation (par Q ou par E avec utilisation du site) de l'aération des informations dans les pages du site.

Concernant la question de l'aération des informations, il peut s'agir d'un besoin en évaluation (« *Mais bon, on peut se le faire confirmer, mais ça on le sait* »), mais les professionnels du Webmastering n'ont pas toute la latitude qu'ils voudraient pour intervenir

sur cet aspect : « Non, ce n'est pas des tensions, c'est des enjeux importants et qu'il y a vraiment des choses à dire... je sais que c'est un des problèmes sur la page d'accueil. Pour l'instant, au niveau de l'information, il y a une telle charge que l'aération n'est pas toujours très aérée. [...] Je pense que c'est une caractéristique qui revient souvent sur les sites marchands » nous explique le responsable de la charte graphique. Si un éventuel travail était possible et réalisé concernant ce point, les professionnels du Webmastering seraient alors certainement preneurs de résultats qui leur fournissent un retour d'appréciation de leur action de conception en termes d'aération.

3.2.4.5. Organisation & Mise en valeur

Tableau 46 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Organisation & Mise en valeur graphique »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Organisation & Mise en valeur graphique	L'organisation ou mise en page des informations dans les pages (d'accueil, de présentation de produits, du panier, etc.) est-elle satisfaisante ?	<p>Explication détaillée des deux caractéristiques (ne pas évaluer chacun des éléments, à moins d'une EA) :</p> <ul style="list-style-type: none"> • Organisation/mise en page : structure, regroupement des éléments similaires, emplacement titre/paragraphe, taille de caractères, etc. • Mise en valeur graphique : mise en couleur des titres, « frames » ou carrés de couleurs différentes pour les diverses rubriques, etc. <p>Un mauvais résultat peut dépendre des éléments précédents.</p>
	La mise en valeur graphique des informations est-elle satisfaisante ?	<p>Appréciation générale de ces deux dimensions concernant les informations (par ordre d'importance) :</p> <ul style="list-style-type: none"> • de la page d'accueil (rubriques, services, titres, slogans, etc.), • des informations importantes sur cette page (cf. catégorie « Page d'accueil ») et sur les autres pages, • des pages d'accueil de chaque famille de produits, • du panier et formulaires de commande, • du « compte perso » (lien vers la page et page avec les services personnalisés), • des listes de résultats à l'issue d'une recherche de produits (cf. aussi « Faciliter à trouver un produit - Liste de résultats »), • des pages de présentation des produits (fiche-produit, en particulier frais d'envoi & disponibilité), • des pages d'informations sur les conditions (de vente, de livraison, de sécurité, les aides à l'utilisation, etc.).

Le responsable de la charte graphique valide cet ensemble. Les éléments de définition de l'organisation/mise en page et de la mise en valeur graphique sont issus des entretiens utilisateurs et de la littérature du domaine (Leulier *et al.*, 1998; Nielsen, 2000; Nogier, 2002).

Par ailleurs, nous avons rassemblé, avec cet ensemble, l'élément « Visibilité des informations importantes » de la même catégorie « Look & co » (présent dans l'ébauche de la boîte à outils, cf. comparaison de l'ancienne et de la nouvelle boîte à outils en annexes du chapitre 6). Nous voulions insister sur la visibilité des informations importantes de la page d'accueil mais les deux éléments sont finalement assez redondants. Le professionnel de la veille-marketing nous le fait remarquer.

Nous avons listé les pages pouvant être concernées par cette évaluation. Ces dernières sont issues de :

- l'énumération déjà énoncée par le responsable de la charte graphique dans l'élément « Reconnaissance visuelle » : la page d'accueil, les pages d'accueil de chaque famille de produits, les outils de la recherche de produits, le panier, les « pages perso »,
- ainsi que des éléments détectés lors de nos entretiens utilisateurs : les informations importantes de la page d'accueil (cf. aussi catégorie « Page d'accueil », rubrique « Réassurance »), les listes de résultats issues d'une recherche (cf. aussi catégorie « Facilité à trouver un produit », rubrique « Liste de résultats »), les pages de présentation des produits (fiche-produit) et les pages d'informations diverses (conditions de vente, aides, etc.).

3.2.4.6. Boutons d'action

Tableau 47 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Boutons d'action »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Boutons d'action (visibilité / emplacements)	Les internautes voient-ils facilement les boutons d'action et les liens ?	<p>Appréciation de leur facilité à voir les boutons d'action et les liens (par Q ou par E avec scénario d'utilisation). Surtout concernant la recherche de produits et le processus d'achat depuis le panier.</p> <p>EA pour comprendre comment les internautes se repèrent dans le site (plutôt par E avec scénario d'utilisation, éventuellement suivis de FG ou par Q ciblé sur les questions de navigation & co). Questions relatives :</p> <ul style="list-style-type: none"> • aux aspects graphiques des boutons et des liens, • à la compréhension de leur code graphique (liens non/déjà visités), • aux emplacements (logiques, constants à travers les pages), • à la formulation de leurs libellés (le libellé correspond-il à ce que l'utilisateur s'attend à obtenir en arrivant sur la page ? l'ajout d'un pop-up sur le lien est-il nécessaire et bien formulé ?), • à la cohérence globale de l'enchaînement des actions, • à leur utilité (ex. boutons pour passer d'une fiche-produit à une autre). <p>Liste des éléments importants : liens, barres de navigation, boutons d'accès au suivi de commande, de mise en panier, d'achat, etc. (cf. aussi « Faciliter à commander un produit »).</p>

Selon le responsable de la charte graphique, il est important de vérifier si les internautes voient bien les liens et les boutons d'action, avec un souci particulier concernant les boutons pour commander et mettre des produits dans le panier. Ce souci est corroboré par les travaux de Tilson *et al.* (1998). Le professionnel de la veille-marketing confirme aussi ce point, ainsi que la plupart des précédents éléments de toute la catégorie « Look & co ».

Nous avons aussi proposé une évaluation approfondie qui porte sur des éléments issus de connaissances ergonomiques concernant les sites Internet et leurs liens (Kim & Yoo, 2000; Leulier *et al.*, 1998; Nielsen, 2000; Nielsen *et al.*, 2001; Nogier, 2002; Yoo & Kim, 2000) : les aspects graphiques des boutons d'action ou des liens ; la compréhension du changement de couleur des liens (non/déjà visités), leurs emplacements ; la formulation des libellés des liens, des boutons d'action et des pop-ups sur ces derniers ; la cohérence globale de l'enchaînement des actions (ex. erreur de compréhension d'un libellé en raison de la signification pour le sujet de l'action précédente) ; et enfin l'utilité de certains liens ou

boutons d'action, notamment ceux qui permettent de passer d'une fiche-produit à la fiche du produit voisin ou de retourner directement à la page d'accueil (*add-on links* qui s'ajoutent à la structure de navigation basique).

Le professionnel de la veille-marketing valide aussi toute l'évaluation approfondie proposée comme étant un besoin en évaluation. Pour cela, nous avons proposé soit une évaluation par questionnaire ciblée exclusivement sur quelques uns des précédents éléments, soit par entretiens individuels avec utilisation du site marchand (observation des comportements des internautes), éventuellement suivis de *focus groups* (étant donné l'intérêt de combiner les deux techniques pour ce genre d'évaluation, cf. chapitre 3). Ce même professionnel fait, en effet, remarquer que tous ces éléments ne pourront pas faire l'objet d'un seul questionnaire d'évaluation : « *C'est trop, c'est trop fouillé. ... parce que s'il est trop questionné, il va en avoir marre. [...] donc il faut surtout pas faire un questionnaire très lourd qui... pour examiner tout ça d'un coup. Il y a plusieurs solutions possibles : soit c'est des bouts de questionnaires, soit c'est un focus group* ».

3.2.4.7. Apparence des courriels

Tableau 48 : Outil d'aide à la conception d'évaluations de sites Internet concernant le besoin en évaluation « Apparence des courriels »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Apparence des courriels	D'un point de vue graphique, les internautes sont-ils satisfaits des courriels ? Ces courriels ont-ils l'impact visuel souhaité ?	Etude complète pour les différents types de courriels plus ou moins personnalisés (Newsletters, offres personnalisées, etc.) sur les dimensions (au choix) : <ul style="list-style-type: none"> • Reconnaissance visuelle (+ évaluation de la réception des courriels : reconnaissance quant à la mémorisation de l'émetteur, de la marque des produits, etc.), • Attractivité, • Graphismes, • Aération, • Organisation & Mise en valeur graphique, • Visibilité /emplacements des boutons d'action. Etude concernant le contenu commercial proposé dans les courriers, cf. catégorie « Netmarketing ».

Dans la catégorie « Look & co », le professionnel de la veille-marketing nous invite vivement à créer une rubrique, à part, pour tout ce qui est non « *pas dans un contexte de consultation mais dans un contexte de réception d'informations* », donc de réception de courriels (« *messages électroniques, c'est quand même un monde en soi* »). Les deux professionnels du Webmastering confirment cela. Le responsable de la charte graphique trouve cela d'un grand intérêt, car il dit avoir très peu d'informations concernant comment

sont perçues les « Newsletters », d'un point de vue graphique. Mais cela doit effectivement constituer une évaluation, à part, et tout ne doit pas être remis en question au même moment (« *actuellement, on remet pas en question les « Newsletters ». Surtout pas tout remettre en question au même moment !* »).

Pour l'évaluation de l'apparence des courriels, nous avons proposé aux professionnels de se référer à toutes les rubriques développées précédemment dans la catégorie « Look & co ». De plus, le professionnel du marketing réagit à la rubrique « Réception des messages publicitaires » (cf. catégorie « Notoriété ») en nous précisant qu'il réalise lui-même ce genre d'évaluation surtout pour les « Newsletters ». Nous avons ainsi proposé d'orienter l'évaluation vers la reconnaissance ou mémorisation de l'émetteur (la marque de produits) des « Newsletters » ou autres courriels.

3.2.4.8. Apparence des « Cross-Sellings »¹⁴⁷ et des Opérations spéciales

Tableau 49 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Apparence des "Cross-Sellings" et des Opérations spéciales »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Apparence des « Cross-Sellings » et des Opérations spéciales	Les internautes les voient-ils facilement sur le site ? Apprécient-ils leurs aspects visuels ?	Etude complète sur les aspects visuels de ces éléments : <ul style="list-style-type: none"> • Attractivité • Graphismes • Organisation & Mise en valeur graphique • Boutons d'action pour la visibilité des liens d'accès aux « Cross-Sellings » et aux Opérations spéciales, dans le site. Etude concernant le contenu commercial proposé dans les offres, cf. catégorie « Netmarketing ».

Le responsable de la charte graphique nous suggère de prévoir également une évaluation des aspects graphiques concernant les opérations spéciales, dont il explique la signification : « *Donc, les opérations spéciales, soit de prix ou de boutique. Ce qu'on appelle boutique transversale. J'ai un Harry Potter, je vais proposer le livre, le CD, le DVD, la petite figurine. Et donc, à ce propos là, je vais faire ce qu'on appelle une boutique, c'est-à-dire une page avec ambiance* ». Il précise aussi ce qu'il est nécessaire d'évaluer : « *Si les gens les voient à partir de bandeaux ou est-ce qu'ils les voient à partir de la page d'accueil ? Ou est-ce qu'ils les voient à partir des pages d'accueil des univers produits, ce qu'on appelle les H.P. ? Est-ce que, pour eux, est-ce que ça leur apporte quelque chose d'avoir du graphisme ? Ca fait partie de 'attractif – pas attractif', vous voyez toutes les questions que vous m'avez posées au début* ». Pour cette évaluation, nous avons proposé aux professionnels de consulter la

précédente rubrique « Attractivité » et aussi les rubriques « Graphismes », « Organisation & Mise en valeur graphique », « Boutons d'action » (pour la visibilité des liens d'accès aux « Cross-Sellings » et aux Opérations spéciales).

De même, le responsable de la charte graphique exprime le besoin d'obtenir des résultats concernant l'apparence des « Cross-sellings ». De plus, les professionnels du marketing sont particulièrement soucieux du « Cross-Selling » ; au même titre que les courriels, les opérations spéciales et la publicité en ligne, bref tout ce qui va inciter le visiteur à acheter. Cette rubrique est donc étendue aux « Cross-Sellings ».

3.2.4.9. Publicité en ligne

Tableau 50 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Publicité en ligne »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Publicité en ligne	Quelles caractéristiques relatives à la publicité en ligne sont les mieux tolérées ?	EA par Q ou par E avec scénario d'utilisation* de différents sites marchands, voire de pages Internet conçues exprès pour l'évaluation du niveau d'acceptation des internautes concernant différentes caractéristiques : <ul style="list-style-type: none"> • format : bannières ou en superposition (« supersticiels » ou « pop-ups ») ou entre les pages (« intersticiels »), • emplacements des publicités dans la page, • espace occupé dans la page, • couleurs, animations graphiques dynamiques, etc. * correspondant à des tâches réalistes d'utilisation : tâche de recherche d'informations ou de produits avec ou sans achat en ligne selon le souhait de l'internaute, tâche de lecture ou d'utilisation explicite de l'élément publicitaire, etc. Cf. aussi « Netmarketing - Promotion "online" des produits »

Cet élément est sujet à controverse. Le responsable du Webmastering ne saisit pas du tout l'intérêt de déterminer quels aspects de la publicité sont les plus mal perçus par les internautes, même s'il est parfaitement conscient de la gêne occasionnée par cette dernière (y compris sur lui-même) : « ... enfin, entre guillemets, on ne va pas acheter des millions, des millions, hein, de bannières en ayant la logique de dire : faisons en sorte qu'on ne la voit pas trop et qu'on ne gêne pas trop les internautes. Non, ce qu'on veut c'est qu'ils cliquent et qu'ils viennent chez nous, c'est pour ça qu'on fait de la publicité ». Après quelques explications du fait que les publicités en ligne constituent l'une des principales frustrations des internautes (Carton, 2002)¹⁴⁸, nous recueillons les commentaires suivants : « Là on a un effet pervers : c'est-à-dire qu'on fait de la publicité qui exaspère, ce qu'on veut faire c'est de la publicité efficace... C'est comme dans ma boîte aux lettres, je ne veux plus de prospectus,

non je voudrais qu'on me mette un prospectus qui va m'intéresser, donc c'est plus ça le ... [...] *Aujourd'hui ils inventent beaucoup de nouveaux formats effectivement pour essayer de trouver des alternatives parce qu'on s'aperçoit que la publicité exaspère beaucoup l'internaute. C'est-à-dire qu'on crée l'effet inverse de ce qu'on voudrait avoir. C'est-à-dire qu'au lieu d'être attractif, on est exaspérant.* » Un phénomène d'« *ad blindness* » a même été observé : les internautes ne regardent pas les publicités pour ne pas être gênés et, de ce fait, ne les voient pas du tout, même quand elles pourraient leur être utiles (Nielsen *et al.*, 2001; Spool *et al.*, 1999).

Enfin, le cœur du problème est peut-être atteint, quand le professionnel finit par nous confier : « *Si on demande aux internautes : comment faire en sorte pour plus que vous ne soyez gênés, ils vont dire... Vous nous enlevez tout !* ». Nous lui expliquons alors qu'il existe diverses techniques en ergonomie permettant d'obtenir des informations assez fouillées sur la perception de divers éléments d'interface, y compris ceux qui sont très contestés. Nous pensons notamment, comme nous l'avons proposé dans la boîte à outils, à une mise en situation du sujet avec utilisation comparée de plusieurs sites marchands existants, voire de plusieurs pages Internet conçues de façon à tester des caractéristiques précises des différentes publicités existantes (format de la publicité : bannière, pop-up, empiètement dans la page ; niveau d'animation dans la publicité ; etc.). Il s'agit là, bien entendu, d'une évaluation assez poussée. Elle est destinée à produire des connaissances sur les aspects les plus néfastes pour proposer des recommandations qui permettent d'en amoindrir les effets. Après toutes ces précisions, le professionnel semble preneur de tels résultats : « *Si, si. Ah, si, si. Oui, non, ça correspond à un besoin.* »

Cependant, la production de ces résultats demeure incertaine, à l'échelle d'une entreprise marchande. Par exemple, il nous semble que l'entreprise Sitemarchand.com n'aurait pas les moyens de financer une telle étude de fond (ils affirment avoir dû déjà abandonner plusieurs activités de communication/animation insuffisamment lucratives). De plus, cette logique ergonomique ne cadre pas avec leur logique commerciale comme nous l'avons vu et comme le remarquent également Spool *et al.* (1999). De plus, *quelques* internautes de nos entretiens utilisateurs expriment une réelle sympathie par rapport aux publicités en ligne, car elles sont attractives et proposent des offres promotionnelles. De même, environ 20% des internautes (singapouriens) de l'étude de Teo (2002) apprécient de lire les publicités par curiosité. Tout cela peut contribuer à ce que les entreprises marchandes délaissent cet aspect.

D'autre part, cet élément était placé dans la catégorie « Netmarketing » dans l'ébauche de la boîte à outils (cf. annexes du chapitre 6), et le professionnel de la veille-marketing pense donc y trouver une évaluation de l'impact de la promotion en ligne des produits. Il nous fait remarquer qu'il s'agit, dans notre proposition, d'évaluer surtout les aspects graphiques et la gêne occasionnée par la publicité en ligne. Cela a donc davantage sa place dans la catégorie « Look & co » : « *Cela ne paraît pas être pertinent dans le paragraphe là. [...] Mais est-ce que ça n'irait pas finalement dans la catégorie look-présentation ? Je ne sais pas... Ah, look peut-être... Look. [...] ou alors dans le look comme tu dis* ». Cet élément a été déplacé vers cette catégorie (cf. comparaison de l'ancienne et de la nouvelle boîte à outils en annexes du chapitre 6).

3.2.5. Catégorie « Offre & Positionnement »

3.2.5.1. Positionnement par rapport à la concurrence

Tableau 51 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Positionnement par rapport à la concurrence »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Positionnement par rapport à la concurrence	Pour les internautes, le site correspond à quel positionnement général par rapport à la concurrence (site « pas cher », site sur lequel on trouve tout, etc.) ?	<p>Etude par Q. Appréciation comparative du site avec les sites (ou les magasins traditionnels) de la concurrence :</p> <p>Par rapport à la concurrence, vous considérez que Sitemarchand.com est :</p> <ul style="list-style-type: none"> • un site pas cher, • un service sérieux, • un site peu recommandable, • un de vos sites favoris, • un site où on trouve tout, • une offre de très haute qualité, • un site dynamique, proposant toujours les derniers produits d'actualité, etc. <p>Puis, comment considérez-vous le site A (de la concurrence) : comme un site pas cher, ...</p> <p>Et, comment considérez-vous le site B (de la concurrence) : comme un site pas cher, ...</p> <p>Etc.</p> <p>EA qualitative par FG (pour élaborer le questionnaire) : comprendre les représentations des internautes quant au positionnement du site. Possibilité de test projectif en FG : si vous deviez comparer le site à un animal, de quel animal s'agirait-il ? ou si vous deviez le comparer à une planète, laquelle vous viendrez à l'esprit ?</p> <p>Rq : compléter par des études de type « benchmark » des offres de la concurrence.</p>

Le professionnel de la veille-marketing nous a incité, lors des précédents entretiens (présentés au chapitre 4), à créer cet élément « Positionnement par rapport à la concurrence », puisque les internautes font des comparaisons entre tout site et tout magasin traditionnel. Nous pensions qu'il s'agissait d'évaluer la satisfaction des cyber-consommateurs concernant le choix de produits et leurs prix par rapport à ceux proposés par les commerces (électroniques ou non) de la concurrence.

Cependant le professionnel nous explique qu'il s'agit, en fait, d'évaluer quelque chose de plus global : la perception des internautes quant au positionnement du site par rapport à la concurrence. « Perception globale. Est-ce que c'est un service sérieux ? Est-ce que c'est un

service où on trouve tout ce qu'on veut ? [...] Parce qu'il peut y avoir une volonté de l'entreprise de se positionner sur tel secteur, sur telle... une façon d'aborder le métier et c'est important de savoir comment c'est perçu. [...] Et [le site A] est perçu comme quoi, et [le site B] est perçu comme quoi, et [le Sitemarchand.com] est perçu comme quoi... est-ce qu'il y a une différence ? Et ça, c'est un élément fondamental des sites de commerce électronique. C'est fondamental, mais ça résume un peu plusieurs questions. Y a ça. Il y a l'image de marque, il y a tout ce qu'on a vu tout à l'heure sur la page d'accueil, c'est-à-dire : est-ce que ça nous concerne, est-ce que ça ne nous concerne pas ? Tout ça, ce sont des éléments qui déclinent le positionnement qui a été déterminé ». Le propos du professionnel est proche du contenu de la rubrique « Image de marque » dans la catégorie « Notoriété ». Néanmoins, les éléments énoncés ici sont à différencier, selon lui, du contenu de cette rubrique : « Mais ça, c'est une mesure de l'image de marque, c'est autre chose qu'une perception du positionnement. [...] C'est très proche quand même. Mais il n'empêche, que c'est... Oui. Quelque chose de plus particulier... »

Ainsi, l'élément « Positionnement par rapport à la concurrence » devient une évaluation comparative de plusieurs sites concurrents, interrogeant l'impression générale des internautes quant au « positionnement » de ces sites (Par rapport à la concurrence, vous considérez que Sitemarchand.com est : un site pas cher, un service sérieux,... ?). Les différents items suggérés pour le questionnaire sont issus du questionnaire d'évaluation Noël 2002 et des propos du professionnel de la veille-marketing.

Nous avons également indiqué une évaluation qualitative soit pour produire directement des résultats subjectifs sur le positionnement du site, soit pour aider à la construction d'un questionnaire d'une évaluation à grande échelle. En effet, les professionnels du (de la veille-) marketing nous affirment que des évaluations qualitatives par *focus groups* ont déjà été réalisées. Le professionnel de la veille-marketing nous précise même que des tests projectifs ont été utilisés pour que les internautes imaginent à quel animal ou à quelle planète le site correspond. Les résultats obtenus étaient très utiles pour comprendre les différentes représentations des cyber-consommateurs concernant le positionnement du Sitemarchand.com.

Enfin, ces données subjectives peuvent être avantageusement complétées par des études de type « benchmark » comparant l'offre du site à étudier avec celle de la concurrence (qualité, marque des produits, variété, prix, etc.).

Etant donné le caractère général et qualitatif de cette évaluation, nous avons placé « Positionnement par rapport à la concurrence » en premier dans cette catégorie (il était à la fin de cette catégorie dans l'ébauche de la boîte à outils, cf. annexes du chapitre 6). Nous avons également ajouté une nouvelle évaluation globale : les motifs d'achat.

3.2.5.2. Les motifs d'achats

Tableau 52 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Les motifs d'achats »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Les motifs d'achats	<p>Quel achat l'internaute réalise-t-il habituellement sur le site (produits rares, des promotions, etc.) ?</p> <p>Avec quelle intention l'internaute vient-il acheter sur le site ?</p> <p>Enquête de fidélisation : Quel est l'élément qui ferait commander l'internaute plus souvent ?</p>	<p>Etude par Q : Vous achetez sur le site parce que vous y trouvez.... ?</p> <p>Liste de réponses possibles :</p> <ul style="list-style-type: none"> • des produits rares, • des pré-commandes, • des nouveautés, • des imports, • des promotions/bonnes affaires • des idées cadeau, etc. <p>Etude par Q : vous arrive-t-il de venir sur le site sans idée d'achat, de réaliser des achats imprévus (toujours, souvent, occasionnellement, etc.) ?</p> <p>Etude par Q. Qu'est-ce qui vous ferait revenir acheter sur le site ?</p> <p>Liste de réponses possibles :</p> <ul style="list-style-type: none"> • l'étendue des produits/services, • la qualité des produits, etc. • les promotions, • les nouveautés, • les informations sur les produits (voire aussi la catégorie « Description des produits »).

Cet élément est assez proche du précédent. Il doit permettre de questionner, de façon plus globale, les motifs d'achat des internautes sur le site. Il nous a été dicté par le questionnaire d'évaluation de Noël 2002 (second questionnaire obtenu) et par le professionnel du marketing. Ce dernier nous explique que la seconde partie de l'évaluation proposée correspond à une enquête de fidélisation : qu'est-ce qui vous ferait revenir acheter sur le site ? Le choix de réponses possibles concernant l'enquête de fidélisation est conforté par divers travaux (Jarvenpaa & Todd, 1997; Jarvenpaa *et al.*, 2000; Konradt *et al.*, 2003; Lohse & Spiller, 1998a, 1998b) tandis que le choix de réponses possibles pour la première partie de

l'évaluation (quel achat l'internaute réalise-t-il habituellement ?) est issu du questionnaire Noël 2002.

3.2.5.3. Choix de produits – Prix – Promotions & Nouveautés

Tableau 53 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant les besoins en évaluation « Choix de produits – Prix – Promotions & Nouveautés »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Choix de produits / services	Les internautes sont-ils satisfaits du choix de produits/services (en termes de variété et de profondeur) ? De la qualité des produits/services ? Des marques proposées ? De l'originalité des produits/services et de leur quantité ?	En fonction des orientations/positionnements choisis par le site marchand, appréciation (plutôt par Q) : <ul style="list-style-type: none"> • de la variété du choix : <ul style="list-style-type: none"> ○ du nombre d'univers de produits différents (ex. CD, Livres, Musique, Vidéo, Hifi, etc.), ○ de l'étendue du choix pour chaque univers de produits différents, • de la profondeur du choix, c'est-à-dire de l'étendue du choix pour chaque gamme de produit (ex. différents modèles d'appareils photos), • de la qualité des produits/services, • des marques proposées, • intérêt/quantité de produits/services originaux, rares (ex. produits étrangers, petits déjeuners à domicile, bouquets de fleurs choisis au diamètre, etc.), • questionnement sur les préférences des internautes quand à la vente d'offres complémentaires (ex. produits d'occasion, billets de spectacles, etc.).
Prix	Comment les internautes perçoivent-ils le rapport qualité-prix ?	En fonction de la politique de prix de l'entreprise (ex. pénétration du marché par des prix bas). Appréciation de la perception (plutôt par Q) : <ul style="list-style-type: none"> • du rapport qualité-prix pratiqué par le site, • ou plus spécifiquement du bas niveau des prix, • éventuellement vérification que les produits sont perçus moins chers ou, du moins, pas plus chers que ceux du commerce classique.
Promotions & Nouveautés	Les internautes sont-ils satisfaits des propositions de promotions, de nouveautés, de pré-commandes ?	Appréciation (plutôt par Q) de l'intérêt (ex. % de réduction du prix) et de l'étendue du choix : <ul style="list-style-type: none"> • des promotions : offres spéciales, prix dégriffés, etc., • des nouveautés ou pré-commandes : produits nouveaux sur le marché, en avant-première, possibilité de les réserver, etc.

Ces trois éléments complètent les deux précédents, voire sont en partie redondants. Ils permettent néanmoins de détailler tout ce qu'il est possible d'interroger dans la satisfaction des cyber-consommateurs concernant le choix de produits, leurs prix et le choix de promotions et de nouveautés. Ils permettent d'approfondir ce qui est précédemment traité de façon globale.

Les professionnels du (et de la veille-) marketing valident tous les besoins en évaluation détaillés dans l'ébauche, excepté certains détails comme l'évaluation du choix de marques proposées qui concernent moins leur secteur de vente. Néanmoins, nous l'avons conservé, car il peut correspondre à d'autres sites marchands. De plus, plusieurs travaux montrent que les cyber-consommateurs sont attentifs aux marques proposées (Jarvenpaa & Todd, 1997; Kehoe *et al.*, 1997).

3.2.5.4. Avantages clients

Tableau 54 : Outil d'aide à la conception d'évaluations de sites Internet marchands g concernant le besoin en évaluation « Avantages clients »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Avantages clients	Les internautes sont-ils satisfaits des divers avantages clients (ex. réduction 1 ^{er} achat, avantages adhérents) ?	<p>Etude par Q. En fonction des orientations et politiques commerciales du site.</p> <ul style="list-style-type: none"> • Appréciation des avantages clients, • Quel est l'avantage le plus apprécié (parmi la liste ci-dessous) ? • Quel est l'avantage le plus utilisé pour les achats ? <p>Corrélation possible avec des données de la base clients (quantité d'achats par type de bons d'achat identifiés à partir de leur numéro, etc.).</p> <p>Liste à titre indicatif :</p> <ul style="list-style-type: none"> • réductions 1^{er} achat, • bons d'achats (ex. après un certain nombre d'achats), • avantages adhérents (ex. réductions avec la carte privative et accumulation de points permettant de bénéficier d'offres-adhérent).

Cet élément faisait partie de la catégorie « Netmarketing » dans l'ébauche de la boîte à outils, mais il est, en fait, très proche de l'offre (notamment de l'enquête de fidélisation : qu'est-ce qui ferait commander l'internaute plus souvent ?). Il apparaît donc plus judicieux de le placer dans cette catégorie « Offre » (cf. comparaison de l'ancienne et de la nouvelle boîte à outils en annexes du chapitre 6), même si les professionnels n'en font pas part. La catégorie « Netmarketing » n'en est que plus homogène (voir plus loin).

Par ailleurs, le professionnel de la veille-marketing confirme la nécessité d'évaluer la satisfaction des internautes concernant les avantages clients et propose un ajout : « *Bon qu'est-ce qu'on peut évaluer là-dessus ? Est-ce qu'ils sont pertinents ? Est-ce qu'ils sont attirants ? Est-ce que vous les utilisez... ?* [Remarque selon laquelle cela semble constituer beaucoup d'éléments] *Oui mais très vite, ça ne sera pas suffisant. [...] mais tu auras envie de savoir rapidement qu'est-ce qui leur plaît, qu'est-ce qui ne leur plaît pas ? Est-ce que c'est*

les bons "1^{ier} achat" qui leur plaisent ? Est-ce que c'est les bons d'achat qui leur plaisent ? ». L'évaluation généraliste prévue dans l'ébauche se voit précisée : il s'agit de connaître la satisfaction générale des internautes quant aux avantages clients, mais aussi de savoir lequel est le plus apprécié, d'une part, et le plus utilisé, d'autre part, parmi une liste présentant l'ensemble des avantages possibles (cf. liste à titre indicatif dans le tableau ci-dessus).

Il est aussi possible de corrélérer les résultats de l'évaluation subjective à des données objectives (quantité d'achats par type de bons d'achat identifiés à partir de leur numéro, etc.). En effet, selon ce professionnel, il est possible d'exploiter la base de données clients pour cela.

3.2.5.5. Positionnement interne

Tableau 55 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Positionnement interne »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Positionnement interne (même enseigne)	Les internautes sont-ils satisfaits du choix de produits, promotions, nouveautés, au regard de ce qui est proposé dans les magasins de la même enseigne ?	EA qui s'apparente à une évaluation de marché. Pour une entreprise marchande disposant de magasins « réels » et de sites Internet (entreprise « traditionnelle » présente sur le Web) : Appréciation générale (plutôt par Q) : <ul style="list-style-type: none"> • du choix de produits du site / celui du magasin, • du niveau des prix du site / ceux du magasin, • des propositions de promotions / celles du magasin, • des propositions de nouveautés / celles du magasin. Recueil des critiques et souhaits divers.

Les propos du professionnel de la veille-marketing (lors de l'entretien avec le Sitemarchand.com, cf. chapitres 4 et 5) et nos entretiens utilisateurs nous ont conduit à créer cet élément, puisque certains internautes disent être très sensibles au fait qu'un site Internet ait au minimum la même offre de produits que ses propres magasins réels et font des comparaisons entre tout site Internet et ses magasins « réels ».

Cet élément constitue un besoin en évaluation spécifique aux entreprises « traditionnelles » présentes sur le Web puisqu'elles ont des magasins réels au contraire des « pures players »¹⁴⁹.

3.2.6. Catégorie « Description des produits »

3.2.6.1. Utilité des informations-produits

Tableau 56 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Utilité des informations-produits »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Utilité des informations produits	Quelles informations les internautes utilisent-ils ou sont les plus utiles pour leur choix de produits ?	<p>Etude par Q (cf. aussi rubriques suivantes). Appréciation générale : Ces informations sont utiles pour choisir un produit, tout à fait d'accord, plutôt d'accord, etc. Liste d'informations-produits (variables selon le secteur de vente) :</p> <ul style="list-style-type: none"> • un résumé, • une photo, • le délai de disponibilité/livraison, • note donnée par les internautes, • commentaires des internautes, • critique de spécialistes, • biographie de l'auteur/réalisateur, • suggestion d'autres produits, • extrait sonore pour la musique, • etc.

Le questionnaire Noël 2002 nous a permis de bâtir cet élément qui cherche à estimer quelle information est la plus utile aux internautes dans leur choix de produits. Il s'agit d'une évaluation comparative ; le fait d'amener le répondant à comparer les différents éléments d'information permet une évaluation plus juste de chacun. Ces résultats doivent aider les professionnels à prioriser et choisir l'information à retenir, puisqu'il est primordial de savoir quel type d'information doit être retenu dans la description des produits (Kim & Lee, 2002).

3.2.6.2. Caractéristiques descriptives des produits

Tableau 57 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Caractéristiques descriptives des produits »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Caractéristiques descriptives des produits	<p>Les internautes sont-ils satisfaits des caractéristiques décrites concernant les produits ?</p> <p>Quelles sont leurs attentes concernant ce point ?</p>	<p>EA par Q ciblée exclusivement sur ce point ou par E avec utilisation du site, suivis de FG pour confronter les éventuels différents points de vue.</p> <p>Appréciation générale concernant :</p> <ul style="list-style-type: none"> • l'utilité de chaque caractéristique descriptive (plus précise que les éléments dans « Utilité des informations-produits », cf. liste en dessous), • leur complétude (éventuels éléments manquants, informations trop vagues, pas assez détaillées, etc.), • leur compréhension (clarté du vocabulaire utilisé), • la fiabilité des informations (si vous avez acheté au moins un produit, trouvez-vous que les informations données correspondent à votre produit ?). <p>Questionnement ouvert : éléments de description que le site marchand idéal devrait présenter ?</p> <p>Rq : Les utilisateurs ont parfois des difficultés à distinguer le contenu en termes d'informations, d'une part, et le look ou la présentation des informations, d'autre part. Il faut donc être attentif à formuler des questions qui clarifient ces différents éléments.</p> <p>Liste à titre indicatif : nom, marque, fabricant, taille-poids-volume, version, puissance, auteur-compositeur, couleur, reliure, finitions, accessoires, options, origine, etc.</p>

Les professionnels du (et de la veille-) marketing trouvent cet élément trop détaillé pour être soumis par questionnaire. Toutefois, nous l'avons conservé ainsi que tous ses éléments, étant donné le nombre d'études, aussi bien avec des internautes qu'avec des Webmasters, qui ont mis en évidence l'importance des informations données sur les produits (Kim & Lee, 2002; Lee *et al.*, 2000; Liu & Arnett, 2000). Il s'agit néanmoins d'une évaluation approfondie, à ne réaliser qu'en cas d'une assez vaste remise en question, par questionnaire exclusivement dédié à la description des produits ou par entretiens avec utilisation du site, suivis d'éventuels *focus groups*.

Le peu d'attention accordée aux informations qui décrivent les produits par les professionnels du (et de la veille-) marketing s'explique du fait que ce sont les professionnels du service produits qui s'occupent de les recueillir et de les transmettre aux agents de saisie qui les entrent quotidiennement dans la base de données des produits. Nous n'avons pas pu rencontrer ces professionnels qui ont sans doute des besoins en évaluation spécifiques. De plus, les professionnels du (et de la veille-) marketing peuvent avoir tendance à freiner cette

évaluation, du fait de l'ampleur de la remise en question qu'elle pourrait engendrer. Il s'agirait éventuellement de modifier les champs de saisie dans la base de données des produits (et de ce fait, l'activité des agents de saisie). Les professionnels du service produits pourraient aussi être amenés à développer une activité de recherche de nouvelles sources d'informations.

Néanmoins, si une évaluation de cette envergure était envisagée, il serait alors utile d'évaluer :

- l'utilité de chaque caractéristique descriptive existante (ex. nom, marque, fabricant, taille-poids-volume, puissance, auteur-compositeur; ce qui est plus précis que les éléments donnés dans « Utilité des informations-produits », cf. liste en dessous),
- la complétude (pour diagnostiquer les caractéristiques manquantes, les informations trop vagues, etc.),
- la compréhension des caractéristiques descriptives des produits (le vocabulaire utilisé engendre-t-il des incompréhensions ou ambiguïtés),
- et aussi la fiabilité de ces informations (les internautes ayant acheté un produit, trouvent-ils que les informations données correspondent effectivement au produit obtenu ?).

Enfin, il faudrait être attentif, dans la formulation du questionnaire ou du guide d'entretien, à clarifier, d'une part, le contenu en termes d'offres et d'informations et, d'autre part, le look ou la présentation des informations, car les internautes peuvent avoir des difficultés à les distinguer (Kwon *et al.*, 2002).

3.2.6.3. Commentaires & Critiques

Tableau 58 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Commentaires & Critiques »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Commentaires & Critiques	<p>Les internautes sont-ils satisfaits des commentaires-critiques donnés sur les produits ?</p> <p>Quelles sont leurs attentes concernant ce point?</p>	<p>Appréciation générale (plutôt par Q) concernant :</p> <ul style="list-style-type: none"> • l'intérêt comparé des différents types de commentaire/critique cf. aussi « Utilité des informations-produits », • la source des commentaires/critiques : quel type de commentaire vous semble le plus fiable : commentaire du site lui-même, de la presse quotidienne, de spécialistes, de l'éditeur, des internautes ? etc. • leur quantité : l'ensemble des commentaires/critiques sont-ils suffisants pour que vous réalisiez votre choix ? Quels autres types de commentaires/critiques seraient utiles ? <p>Rq : l'éditorial se situe généralement sur les fiches produits, mais aussi sur certains sites, dans des rubriques éditoriales spécifiques, une sélection du mois, etc. si le site le développe (nécessite un service éditorial à part entière).</p>

De même, cet élément semble trop détaillé aux professionnels du (et de la veille-) marketing. Néanmoins, nous l'avons maintenu étant donné l'importance de l'éditorial et des commentaires des experts ou d'autres internautes sur les produits (Kim & Lee, 2002; Nielsen *et al.*, 2001).

Cette évaluation est complémentaire à la rubrique « Utilité des informations-produits », mais est ciblée sur les caractéristiques des commentaires/critiques (en plus de leur utilité). Il s'agit d'évaluer la satisfaction des internautes quant :

- au choix de la source des commentaires/critiques (élément issu du questionnaire Noël 2002),
- ou à leur quantité sur le site (élément issu des entretiens utilisateurs).

3.2.6.4. Illustrations

Tableau 59 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Illustrations »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Illustrations	<p>Les internautes sont-ils satisfaits des illustrations ?</p> <p>Quelles sont leurs attentes concernant ce point?</p>	<p>Appréciation générale (plutôt par Q ou par E avec utilisation du site) concernant :</p> <ul style="list-style-type: none"> • la taille des illustrations (bonne visibilité du produit), • la taille des photos une fois agrandie (si la possibilité existe), • la rapidité et le mode de chargement des photos, • éventuellement, l'intérêt de l'illustration (ex. photo de l'hôtel à réserver plutôt que d'un paysage quelconque) et du support d'illustration (photo ou 3D). <p>Liste dans l'ordre du plus courant et du plus recommandé : photo, vidéo, dessin-schéma, illustration 3D, 3D mobile, etc.</p>

Le responsable de la charte graphique valide cet élément et l'enrichit de la question de la taille des photos : les photos sont-elles assez grandes ou assez claires pour rendre compte du produit correctement (très valable, par exemple, pour les vêtements ou les petits produits technologiques comme la connectique) ? Même question pour la photo agrandie, si le site le propose. Ces illustrations sont, en effet, essentielles (Nielsen, 2000) et les internautes se plaignent de l'impossibilité de visualiser correctement le produit, tout autant sinon davantage que sur les catalogues papier (Jarvenpaa & Todd, 1997).

De plus, le professionnel de la veille-marketing nous suggère d'ajouter ici la question du chargement des photos, alors que nous l'avons intégrée dans la catégorie « Rapidité d'accès – Capacité système ». Il est plus logique d'aborder ce point ici, de façon contextualisée. Il s'agit d'évaluer la rapidité de chargement, ainsi que le mode de chargement des photos, puisque ce dernier influence la tolérance des internautes, comme nous l'avons vu à la rubrique « Rapidité de chargement ».

Enfin, l'intérêt de l'illustration nous a été suggéré par nos entretiens utilisateurs¹⁵⁰ puisque certains des internautes trouvaient qu'il aurait mieux valu avoir la photo de l'hôtel à réserver plutôt que celle d'un paysage standard « façon carte postale » ou désiraient des illustrations 3D pour voir, par exemple, les vêtements « sous toutes leurs coutures ». Il faut néanmoins rester vigilant sur ce point : les internautes ont tendance à demander beaucoup, et cela alourdirait d'autant le poids des pages. En effet, la parcimonie est recommandée concernant les éléments en 3D et rares sont les situations où la 3D présente un réel avantage (Nielsen, 2000).

3.2.6.5. Annonce prix

Tableau 60 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Annonce prix »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Annonce prix	Les internautes sont-ils satisfaits de la façon d'annoncer les prix ?	EA. Appréciation (par Q ou par E avec utilisation du site) du choix des éléments pour détailler le prix ? Etude particulièrement indiquée si les prix sont perçus peu compétitifs par les internautes, alors qu'ils le sont en réalité. Liste d'éléments possibles d'annonce du prix : <ul style="list-style-type: none"> • nos prix / prix habituels, • montant du gain effectué, • prix calculé avec la réduction incluse et non pas seulement -5% à côté, etc.

Cet élément qui complète la rubrique précédente « Prix », trouve également sa source dans les entretiens utilisateurs, en plus d'études comme celle de Kim et Lee (2002). Les internautes se plaignent de certaines lacunes : ils veulent voir mis en évidence le prix calculé avec la réduction annoncée et non pas seulement « -5% » affiché à côté ou encore le gain qu'ils réaliseraient en achetant par Internet.

De plus, cet élément est validé par le professionnel de la veille-marketing. Il nous relate, en effet, une enquête ayant mis en évidence que les prix du Sitemarchand.com étaient mal perçus, alors qu'ils étaient compétitifs par rapport à la concurrence. Ce professionnel est parfaitement conscient des nombreux facteurs d'annonce intervenant dans cette perception des prix. Cet élément constitue une évaluation approfondie par questionnaire ou par entretiens (avec mise en situation d'utilisation du site marchand), à réaliser seulement s'il s'avère que les prix sont mal perçus.

3.2.6.6. Annonce disponibilité & frais d'envoi

Tableau 61 : Outil d'aide à la conception d'évaluations de sites Internet concernant le besoin en évaluation « Annonce disponibilité & frais d'envoi »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Annonce disponibilité & frais d'envoi	Les internautes sont-ils satisfaits de l'annonce concernant les disponibilité et frais d'envoi ?	<ul style="list-style-type: none"> • Appréciation par Q de la satisfaction des internautes quant aux informations fournies sur la disponibilité & frais d'envoi (sur les fiches-produit en particulier et non pas seulement dans les conditions de vente. Si ces éléments figurent sur les fiches-produit, cette évaluation n'est pas nécessaire). • Appréciation par E de la compréhension/clarté de l'annonce (ex. « en stock », « généralement disponible en 24h », etc.). Quelle autre formulation conviendrait mieux ?
Aspects visuels de toutes les informations produits		Cf. catégorie « Look & co ».

Les deux professionnels du (de la veille-) marketing ne sont pas, non plus, très « loquaces » sur les éléments d'annonce de la disponibilité et des frais d'envoi. Nous avons cependant maintenu ce point, étant donné les nombreuses réclamations des internautes lors de nos entretiens utilisateurs et la littérature du domaine qui les soutiennent. Il est notamment dénoncé le fait que l'utilisateur découvre souvent les frais d'envoi au stade ultime de sa commande (Fouchard, 2001; Nielsen *et al.*, 2001). L'absence d'informations *sur chaque fiche-produit*, concernant les frais d'envoi et la disponibilité du produit, est également déplorée. Enfin, la clarté de certaines annonces n'est pas en reste. Par exemple, la formule laconique « généralement disponible en 24h » est assez mal perçue. « En stock » n'est pas non plus très précis. Les internautes préfèrent connaître le délai habituel de réception du produit à domicile (Nielsen *et al.*, 2001).

Nous avons proposé une évaluation par questionnaire pour estimer la satisfaction des internautes quant aux informations fournies sur la disponibilité des produits et leurs frais d'envoi (de préférence en précisant sur chaque fiche-produit) et une évaluation par entretiens pour estimer les problèmes quant à la clarté des annonces.

La visibilité de l'ensemble des informations concernant les produits peut également être évaluée et nous l'avons indiqué aux professionnels en les renvoyant à la catégorie « Look & co ».

3.2.6.7. Mise à jour – « Fraîcheur » des produits

Tableau 62 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Mise à jour – "Fraîcheur" des produits »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Mise à jour – Fraîcheur des produits	Les internautes ont-ils le sentiment que les produits et informations sont renouvelés et mis à jour (suffisamment fréquemment) ?	EA. Dans le cas des sites marchands peu connus ou qui ne réalisent pas de mises à jour quotidiennes (cette évaluation n'a pas de sens pour les sites qui réalisent des mises à jour quotidiennes). Appréciation générale (plutôt par Q) : <ul style="list-style-type: none"> • de la « fraîcheur » des produits et informations sur les produits, • de la pertinence des indicateurs de mise à jour (date sur la page d'accueil, sur chaque fiche-produit si les produits sont peu nombreux ou particulièrement rares, etc.).

A vrai dire, cet élément est réfuté par tous les professionnels, car le Sitemarchand.com est mis deux fois à jour quotidiennement. Même si aucun message n'informe de cela sur le site, interroger ce point ne leur semble pas pertinent. Il leur semble que les internautes ne se soucient pas de cette question : « *Oui, mais le client est-ce qu'il vient assez souvent pour ... [...] Donc quand il sera sur le site, il va aller voir une fois le CD, par exemple [...], si le CD lui plaît, si l'information le décrivant est satisfaisante, il va l'acheter mais il ne va pas revenir dessus après* ». Cela est juste, néanmoins certains internautes se plaignent de l'affichage de produits obsolètes et sont sensibles à l'existence d'une date (sur la page d'accueil) informant de la dernière mise à jour, d'autant plus si l'entreprise marchande « virtuelle » ou « terrestre » est peu connue. Pour une clientèle internationale, il est même suggéré d'afficher la date en précisant s'il s'agit de l'heure locale (Fouchard, 2001).

Ainsi, cette évaluation concerne la perception des internautes quant à la « fraîcheur » des produits et à l'utilité d'indicateurs de mise à jour. Elle est limitée aux sites marchands peu connus et/ou n'ayant pas les moyens de mettre leur site à jour quotidiennement, mais soucieux de vérifier comment leur site est perçu sur ce point.

3.2.6.8. Editorial

Tableau 63 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Editorial »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Editorial (valeur ajoutée)	Quelle valeur ajoutée apporte la présence de l'éditorial ?	EA : évaluation très approfondie et coûteuse. Etude à partir de cyber-consommateurs achetant sur des sites éditoriaux (nécessite de pouvoir recruter de tels acheteurs). Par E individuels spécifiques (E d'explicitation d'une séquence précise relative à la consultation de l'éditorial). Appréciation de l'utilité de l'éditorial pour leur décision d'achat,

Cette évaluation concernant la valeur ajoutée qu'apporte un véritable service éditorial interne (formé de journalistes et/ou de critiques expertes) trouve son origine dans une interrogation du professionnel de la veille-marketing (lors des premiers entretiens libres décrits au chapitre 4). Il nous confirme ce besoin en évaluation, tandis que le professionnel du marketing émet des doutes quant à la rentabilité même d'une telle évaluation.

Il s'agit, en effet, d'une évaluation approfondie et coûteuse qui apporte surtout des résultats qualitatifs. Nous avons proposé d'utiliser l'entretien d'explicitation (réalisé par un ergonome formé à cette technique). L'entretien d'explicitation porterait sur une expérience vécue de choix de produits, suite à la lecture de rubriques éditoriales¹⁵¹. A travers ces entretiens, il s'agit d'estimer l'impact de la lecture de l'éditorial dans leur choix de produits. Pour ce faire, il faut recruter des internautes « lambdas » ayant réalisé plusieurs achats dans les six derniers mois sur un site avec éditorial. Un cabinet ou un service spécialisé dans le recrutement (comme celui du service d'évaluation de FTA) peuvent réaliser cela, s'ils possèdent une base de données de type « panel » précisant plusieurs caractéristiques de participants potentiels.

3.2.7. Catégorie « Netmarketing »

Du fait de son ampleur, cette catégorie est scindée en deux sous catégories : « Esprit communautaire et offres spéciales ou croisées », d'une part, et « "Newsletters", courriels "en push" et "compte perso" », d'autre part.

3.2.7.1. Esprit communautaire et offres spéciales ou croisées

Esprit communautaire

Tableau 64 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Esprit communautaire »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Esprit communautaire	D'une façon générale, les internautes ont-ils le sentiment que l'esprit communautaire (achats et avis des autres internautes, possibilité de voter, etc.) est bien développé ?	Appréciation (plutôt par Q) : <ul style="list-style-type: none"> • du développement de l'esprit communautaire sur le site (bien développé, peu, etc.), • ou si le site n'a pas développé cet aspect, chercher à interroger les attentes des internautes : qu'est-ce qu'ils souhaiteraient particulièrement voir développer sur le site ? • de l'utilité comparée des différentes fonctions (présélections de produits dans les listes de résultats, « Cross-Sellings », vote électronique, rubrique « avis de consommateurs », etc.). Cf. aussi les différentes rubriques ci-dessous.

La première génération de l'Internet, de « *culture anticonsommation, antimonopoles, antimarques* », bref baptisée la cyber-culture, est attachée à une certaine éthique, à l'idée de réseau et de partage de l'information, tandis que la seconde population, plus large, « *ne partage pas nécessairement les mêmes valeurs* » et « *ignore a priori les valeurs fondatrices de l'Internet et du Web* » (Fouchard, 2001). En effet, aux Etats-Unis, le profil de la population des internautes est de plus en plus similaire à celui de la population « nationale », en termes d'âge, de revenu, d'éducation, etc. Ce phénomène semble moins prononcé en Europe (Kehoe *et al.*, 1998a; Lohse *et al.*, 2000).

Toujours est-il que cet aspect communautaire du Web a pu être, en partie, conservé, tout en se traduisant de façon beaucoup plus commerciale. Il s'agit dorénavant des présélections de produits dans les listes de résultat (ex. produits les plus populaires ou les plus vendus), des propositions de « Cross-Sellings », de la possibilité de vote électronique et des rubriques « avis de consommateurs » où il est possible de lire et de déposer des témoignages.

Les professionnels du (et de la veille-) marketing et même le responsable de la charte graphique se montrent très soucieux de ces aspects. Nous avons donc ajouté cette rubrique qui concerne l'esprit communautaire du site dans son ensemble, avant de traiter des évaluations dédiées exclusivement à chaque élément : présélections, « Cross-Sellings », vote électronique, etc.

Dans cette rubrique, il s'agit d'évaluer la satisfaction des internautes quant au développement de l'esprit communautaire sur le site (bien développé, peu, etc.), puis de

chercher à connaître les attentes des internautes en la matière, particulièrement si le site n'a pas encore développé cet aspect. Les propos du responsable de la charte graphique nous éclairent sur ce point : « *Oui mais alors, ça ne suffirait pas de poser juste ça, parce que moi, je connais la réponse. Ils vous diront non, parce qu'on n'a pas développé l'aspect communautaire sur [le Sitemarchand.com]. Mais est-ce que ce serait souhaitable que [le Sitemarchand.com] développe... et comment est-ce que vous aimeriez qu'il y ait un développement des espaces...* ».

Enfin, toujours dans cette rubrique, nous avons proposé l'évaluation de l'utilité comparée des différentes fonctions communautaires (« Cross-Sellings », vote électronique, etc.) qui doit renseigner les professionnels sur les fonctions nécessitant une évaluation plus approfondie.

Les éléments suivants se recoupent en partie avec cette rubrique. Chaque professionnel peut néanmoins choisir ce qu'il veut retenir selon sa situation (relative au site à étudier et au moment de l'évaluation).

Vote électronique & Témoignages

Tableau 65 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Vote électronique & Témoignages »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Vote électronique & Témoignages	<p>Les internautes trouvent-ils utile de pouvoir donner leurs avis concernant le site ou les produits ?</p> <p>Trouvent-ils utiles de lire les avis des autres internautes ?</p>	<p>Etude par Q (qui peut varier selon les fonctionnalités proposées sur le site). Appréciations de :</p> <ul style="list-style-type: none"> • 1/ l'utilité d'une rubrique « avis de consommateurs » (très utile, utile, peu utile, ...), • 2/ l'utilité du vote électronique sur les produits, • 3/ la quantité d'avis d'autres internautes (suffisante ou non). <p>EA par Q ou par E avec utilisation du site. Questionnement concernant 1/ et 2/ :</p> <ul style="list-style-type: none"> • 1/ Les internautes lisent-ils souvent la rubrique « avis de consommateurs » ? Déposent-ils souvent un avis ? Pourquoi, dans quel objectif ? Qu'y manque-t-il ? Qu'est-ce qui pourrait l'améliorer ? • 2/ Consultent-ils souvent les votes sur les produits ? En quoi sont-ils utiles pour eux ? Votent-ils souvent eux-mêmes ? Que manque-t-il concernant cette fonction ? Qu'est-ce qui pourrait l'améliorer ?

Cet élément faisait partie de la catégorie « Interactivité » dans l'ébauche de la boîte à outils (cf. annexes du chapitre 6) qui était scindée en deux sous-catégories : l'interactivité directe et l'interactivité en ligne. Cependant, étant donné les regroupements que font les professionnels, nous avons « rapatrié » cet élément ainsi que « Liens vers des sites

d'informations, des chats, des forums » dans le « Netmarketing ». La catégorie « Interactivité » n'en est que plus homogène puisqu'elle ne concerne plus que l'interactivité directe entre le client et l'entreprise marchande (même si elle peut se faire via courriels), comme nous le verrons plus loin.

Cet élément vise à évaluer la satisfaction des internautes concernant l'utilité du vote électronique, l'utilité de la rubrique « avis de consommateurs » et la quantité de témoignages déposés. Par exemple, on sait que les jeunes internautes, en particulier, sont friands de la possibilité de voter sur un produit, de déposer des témoignages ou de lire les avis de ceux qui ont déjà acheté le produit (Lambert & Legrand, 2001).

De plus, l'existence de ces fonctionnalités fait augmenter les ventes (Lohse & Spiller, 1998b). Les professionnels sont conscients de ces enjeux et approuvent tous ces éléments. Nous avons proposé une évaluation d'abord généraliste, puis une évaluation approfondie, si nécessaire, par questionnaire ou par entretiens avec utilisation du site. Il s'agit alors d'approfondir les représentations et jugements des internautes et de déceler ce qui est lacunaire : consultent-ils ou déposent-ils souvent des commentaires ou des votes sur les produits ? quand ils les consultent, qu'y cherchent-ils et qu'y trouvent-ils exactement ? et surtout, qu'est-ce qui pourrait améliorer ces fonctions ?

Présélections dans les listes de résultats

Tableau 66 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Présélections dans les listes de résultats »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Présélections dans les listes de résultats	Les internautes sont-ils satisfaits du choix de produits dans les présélections en entête des listes de résultats, ainsi que de l'optique choisie pour les définir ?	<p>Appréciation (plutôt par Q) :</p> <ul style="list-style-type: none"> • du choix de produits dans les présélections (les produits proposés dans les présélections vous semblent-ils intéressants ?), • des préférences des internautes concernant l'optique de ces présélections (dans les présélections, préférez-vous obtenir les produits les plus vendus, les plus en vogue – sélection éditoriale ou à partir des votes électroniques, les moins chers, etc. ?). <p>Cf. aussi catégorie « Facilité à trouver les produits » concernant l'aide qu'apportent les présélections pour trouver un produit.</p>

Les présélections¹⁵² en entête des listes de résultats peuvent guider les recherches des internautes et leur donner une idée immédiate de l'offre de produits d'un site. Faites l'expérience d'entrer « guide Paris » dans la recherche par mots-clefs de deux sites différents de produits culturels, vous n'obtiendrez pas les mêmes résultats en présélections et

probablement vous vous ferez ainsi une idée a priori du site. C'est certainement ce que le professionnel de la veille-marketing a voulu signifié quand il dit : « *Oui..... vous concerne ou correspond à votre recherche, quoi. Je ne sais pas comment dire... Que le site vous concerne, c'est dans le sens que on a fait une recherche et puis on essaie de voir en quelques secondes, si ce site correspond à ce qu'on voulait...* ».

Nous avons proposé d'évaluer la satisfaction des internautes quant au choix de produits obtenus dans les présélections, ainsi que leurs préférences concernant l'optique choisie pour ces présélections (produits les plus vendus, les plus en vogue, les moins chers, etc.).

« *Cross-Sellings* »

Tableau 67 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « *Cross-Sellings* »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
« <i>Cross-Sellings</i> »	Les internautes sont-ils satisfaits du contenu dans les propositions de « <i>Cross-Sellings</i> » ?	Appréciation (plutôt par Q) de : <ul style="list-style-type: none"> • l'utilité des offres de « <i>Cross-Sellings</i> » (regardent-ils ou achètent-ils souvent ces offres ? dans quelle proportion), • l'intérêt de chaque offre en termes de produits proposés, selon les optiques choisies par l'entreprise (une offre d'accessoires complémentaires, de produits moins chers, de différentes versions ou capacités d'un produit, de produits qui vont bien ensemble, etc.), • la pertinence des informations d'accompagnement (s'il y en a).

Cet élément est le complémentaire du précédent. Nous avons proposé d'estimer :

- si les internautes utilisent les propositions de « *Cross-Sellings* » pour guider leurs choix de produits,
- leur satisfaction quant au contenu en termes de produits dans les propositions de « *Cross-Sellings* », à ajuster en fonction des optiques choisies par l'entreprise, puisqu'il existe différents principes de « *Cross-Sellings* » (une offre d'accessoires complémentaires, de produits moins chers, etc.),
- et enfin la pertinence des informations qui les accompagnent, quand il y en a.

Le professionnel de la veille-marketing valide tous ces éléments et conseille une évaluation par questionnaire.

« *Présélections dans les listes de résultats* » et « *Cross-Sellings* » ont été inversés ; « *Cross-Sellings* » était en premier dans l'ébauche de la boîte à outils (cf. annexes du chapitre 6). Le professionnel de la veille-marketing a, en effet, des difficultés à identifier ce que représentent les présélections dans les listes de résultats. Or, ces dernières constituent la

première étape d'une recherche de produits par mots-clefs avant de sélectionner un produit et de rencontrer des offres croisées complémentaires. Il nous invite ainsi à inverser l'ordre chronologique des deux.

S'il s'avère que les internautes sont insatisfaits des offres de produits croisées ou dans les présélections dans les listes de résultats, les informaticiens peuvent être sollicités pour réviser et affiner les outils de recueil et d'analyse de données sur les internautes (leur parcours sur le site, les achats qu'ils réalisent, etc.).

Opérations & Boutiques spéciales en ligne

Tableau 68 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Opérations & Boutiques spéciales en ligne »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Opérations & Boutiques spéciales en ligne	Les internautes trouvent-ils utiles les opérations spéciales (page spéciale proposant un assortiment de produits) ?	<p>Appréciation (plutôt par Q) :</p> <ul style="list-style-type: none"> • de l'intérêt des offres spéciales en ligne, • dans le cas où ses offres sont considérées comme peu intéressantes, pourquoi (autres attentes, lesquelles ? question ouverte). <p>Corrélation possible avec les taux d'audience et de transformation à partir de ces pages (estimation du volume de ventes liées à ces offres).</p>

Cet élément n'existait pas dans l'ébauche de la boîte à outils (cf. annexes du chapitre 6). Nous l'avons ajouté étant donné que plusieurs professionnels (Webmastering et marketing) évoquent les boutiques ou opérations spéciales en ligne¹⁵³ sous divers aspects (notamment graphiques, cf. catégorie « Look & co »).

Dans la même optique que tout ce qui précède, il s'agit d'évaluer l'intérêt du contenu de ces offres spéciales. L'évaluation doit aussi permettre de comprendre en quoi ce contenu est éventuellement inapproprié (quelles sont les autres attentes des cyber-consommateurs ?).

De plus, nous avons indiqué la possibilité de corréler les résultats de l'évaluation subjective à des données du taux d'audience (et du taux de transformation par extrapolation) comme indicateurs de réussite ou d'échec de ces opérations. Le professionnel du marketing nous fait, en effet, remarquer qu'il peut estimer l'impact de ses opérations en consultant l'évolution du taux d'audience à la suite de leur mise en œuvre. « La base d'audience, vous ne vous en occupez pas. Elle sert à quoi cette...? C'est un baromètre ? *Par exemple..., par exemple, pour reconnaître dans l'audience du site, l'audience qui est faite par mes opérations* ».

Promotion « online » des produits

Tableau 69 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Promotion "online" des produits »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Promotion « online » des produits	<p>Les opérations publicitaires ont-elles une influence sur les performances du site ?</p> <p>Les publicités sont-elles intéressantes pour les internautes ? Sinon, pourquoi ?</p>	<p>Etude par Q en ligne à la suite d'un achat réalisé par un internaute (rq : cette évaluation n'interroge que les internautes ayant acheté et n'est pas représentative de la totalité des visiteurs/clients du site).</p> <p>Questionnement : par quel moyen a-t-il découvert le produit :</p> <ul style="list-style-type: none"> • bouche à oreille, • un bon d'achat ou une offre personnalisée, • une publicité en ligne, • toute publicité hors Web, etc. <p>Corrélation avec le taux d'audience des publicités du site (nombre de clics par publicité).</p> <p>EA sur la publicité en ligne (par Q ou par FG).</p> <p>Appréciation générale :</p> <ul style="list-style-type: none"> • selon différentes dimensions (gênante, attirante, amusante, utile, répulsive, etc.), • éventuellement du thème des publicités en lien avec le site (ex. refus de publicités pornographiques sur un site de vin), • questionnement ouvert pour mettre en évidence certaines déceptions (ex. les publicités qui font la promotion d'un produit et qui renvoient à la page d'accueil du site au lieu de la fiche du produit). <p>Cf. aussi « Look & co – Publicité en ligne »</p>
Aspects visuels des « Cross-Sellings », des Opérations spéciales et de la publicité		Cf. catégorie « Look & co »

Cet élément est issu des propos du professionnel de la veille-marketing commentant la rubrique « Publicité en ligne » de l'ébauche de la boîte à outils et s'attendant à y trouver un élément de l'ordre de la promotion « online » des produits vendus sur le site : « *je pensais que tu évaluais ici la publicité que fait le site pour ses produits* ».

Etant donné qu'il s'agit là de questionner comment la découverte d'un produit s'est faite, nous avons réutilisé l'idée de l'enquête en ligne¹⁵⁴, puisque cela semble indiqué : « *Mais ça c'est possible aussi, c'est-à-dire qu'on peut très bien faire une enquête après achat. A la fin de l'enquête d'achat, tiens vous avez acheté ça, comment ça se fait ? Pourquoi avez-vous acheté ça ?* ».

Suite aux propos du même professionnel, nous avons aussi indiqué la possibilité de corréler les résultats de l'évaluation subjective à des données objectives, c'est-à-dire des

mesures spécifiques de l'audience des publicités en ligne : nombre de clics effectués, nombre de « clickthroughs » où la page entière contenant la publicité a été affichée, etc. (Fouchard, 2001).

En dehors des éléments évoqués par les professionnels, nous avons ajouté une évaluation approfondie par entretiens permettant, à nouveau, de comprendre ce qui est mal perçu par les internautes concernant les publicités en ligne. Seulement cette fois-ci, il ne s'agit pas de rechercher les aspects graphiques gênants, mais de recueillir des données relatives à :

- la satisfaction générale des internautes au regard de diverses dimensions¹⁵⁵ (publicité gênante, attirante, amusante, utile, répulsive, etc.),
- leur gêne en rapport avec les thèmes des publicités (ex. publicités pornographiques sur un site de vin),
- à leurs insatisfactions quant au fonctionnement des publicités en ligne (par exemple, un tiers environ des internautes quittent immédiatement un site, lorsqu'un de ses bandeaux publicitaires les mène vers la page d'accueil d'une entreprise et non vers la fiche du produit promu (Nielsen, 2000).

Enfin, nous avons indiqué aux professionnels la catégorie « Look & co » pour traiter des « Aspects visuels des "Cross-Sellings", des opérations spéciales et de la publicité ».

Liens vers des sites d'informations, des chats, des forums

Tableau 70 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Liens vers des sites d'informations, des chats, des forums »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Liens vers des sites d'informations, des chats, des forums	Les internautes sont-ils satisfaits des possibilités d'accès à des informations ou produits complémentaires ou à des discussions en ligne ?	<p>Selon les possibilités de communication développées par le site et selon les produits vendus (ex. achat très impliquant comme l'achat d'une piscine), EA par E ou par Q. Appréciations de :</p> <ul style="list-style-type: none"> • la pertinence des informations vers lesquelles les internautes sont envoyés (ex. liens vers des sites de comparatifs de prix, vers la biographie d'un auteur, la description de pays), • la pertinence des produits complémentaires disponibles à partir d'un lien vers un autre site marchand (ex. livres techniques sur la plongée sous-marine à partir d'un site vendant des produits de ce secteur), • l'intérêt des discussions en ligne (en lien avec l'offre du site et ses orientations), • leur quantité suffisante.

Les professionnels du marketing comme du Webmastering, sont tous catégoriques : il ne faut mettre aucun lien qui soit susceptible de faire partir l'internaute.

Pourtant, ce n'est pas l'avis de plusieurs auteurs et praticiens du domaine. En effet, les liens externes vers des informations complémentaires ne coûtent presque rien et permettent à un site Internet de profiter de la richesse du contenu de millions d'autres sites (Nielsen, 2000). De plus, les liens d'informations complémentaires, à partir de sites Internet marchands notamment, sont très appréciés (ex. diverses sources d'informations sur le camping à partir d'un site vendant des équipements de loisir de plein air ; Fouchard, 2001; Jarvenpaa & Todd, 1997). La présence de liens, non seulement vers des sites d'informations, mais aussi vers d'autres sites marchands est même préconisée (Fouchard, 2001). Cela peut s'établir entre sites marchands du même groupe commercial (ex. différentes entreprises marchandes de la filiale de France Télécom), mais il peut aussi y avoir des accords commerciaux entre commerces électroniques de secteurs de vente complémentaires.

Le responsable du Webmastering partage, en fait, l'avis selon lequel les liens externes constituent une valeur ajoutée pour un site marchand, mais un obstacle majeur l'empêche de se pencher sur la question, la rentabilité commerciale immédiate ! : « ça crédibilise le site, hein si on donne l'information qu'il [l'internaute] veut et des liens vers des sites très intéressants et en adéquation avec ses centres d'intérêt ... Mais c'est vrai qu'on est tellement dans une logique de rentabilité immédiate, en ce moment, sur le Web. [...] Quand on sait qu'un bon internaute ne va acheter que 2 fois et demi par an. Moi, je ne peux pas me

permettre de le laisser partir 6 mois et de le laisser revenir. Parce que, dans 6 mois, je ne suis pas sur qu'on sera encore là, hein. ».

De même, l'existence de chats et de forums permet de faire communiquer les internautes entre eux et augmente leur intérêt dans le site (Fouchard, 2001; Lambert & Legrand, 2001; Nielsen *et al.*, 2001; Teo *et al.*, 2003). Cependant, une fois de plus, concrètement le problème est le manque de moyens alloués à cette activité : « Mais, par contre, l'autre fois [entretien un an auparavant], il me semble que vous m'aviez parlé aussi d'un forum lié à la fête de la musique. *Mais ça on ne le fait plus du tout, par contre. Parce qu'ils se sont aperçus que ça nous demandait beaucoup de moyens financiers, beaucoup de ressources intérieures pour gérer ou mettre en place et que derrière il n'y avait absolument pas de rentabilité ».*

Toute cette discussion pour expliquer pourquoi nous avons gardé cet élément, du moins sous la forme d'une évaluation approfondie, qui cherche à évaluer la satisfaction des internautes concernant :

- la pertinence des informations ou des produits complémentaires, accessibles par des liens renvoyant vers d'autres sites marchands ou non,
- l'intérêt des discussions en ligne et leur quantité suffisante.

Des entreprises marchandes développant davantage ces aspects informationnels et interactifs peuvent être intéressées par ces évaluations.

3.2.7.2. « Newsletters », courriels « en push » et « compte perso »

Courriels en « push » & « Newsletters » ou offres personnalisées

Tableau 71 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant les besoins en évaluation « Courriels en "push" & "Newsletters" ou offres personnalisées »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Courriels en « push » (<i>Opt-in / Opt-out</i>)	Les internautes sont-ils satisfaits du contenu des courriels en « push » ?	Rq : courriels publicitaires ou d'offres diverses (personnalisées ou non) constituées à partir de données recueillies sur les internautes (sur la navigation des visiteurs, sur les commandes effectuées en ligne par les clients, lors de leurs appels au service clients, etc.). Appréciation générale (plutôt par Q) : <ul style="list-style-type: none"> de l'offre en termes de produits proposés, de toutes les informations, slogans, etc. accompagnant l'offre, de la quantité de ces envois. Recueil plus libre de souhaits et suggestions diverses (question ouverte).
Newsletters non personnalisées	Les internautes sont-ils satisfaits du contenu proposé dans les Newsletters classiques ?	Rq : Newsletters type lettres d'actualités auxquelles les internautes choisissent de s'abonner (mais qui sont identiques pour tous) proposant diverses offres, annonçant des événements commerciaux, etc. Appréciation générale (plutôt par Q) : <ul style="list-style-type: none"> de l'offre en termes de produits proposés, de toutes les informations, slogans, etc. accompagnant l'offre. Recueil plus libre de souhaits et suggestions diverses (question ouverte).
Newsletters personnalisées	Les internautes du site sont-ils satisfaits du contenu des Newsletters ou autres offres personnalisées ?	Rq : les Newsletters ou autres offres personnalisées sont réalisées pour chaque profil client spécifié volontairement par ce dernier (ex. correspondant à leurs centres d'intérêt, ex. rock et jazz en musique, polar et cuisine en littérature). Appréciation générale (plutôt par Q) : <ul style="list-style-type: none"> de l'offre en termes de produits proposés, de toutes les informations, slogans, etc. dans la Newsletter. Recueil plus libre de souhaits et suggestions diverses (question ouverte).
Aspects visuels de tout courriel		Cf. catégorie « Look & co »

Voici un autre sous-ensemble de la catégorie « Netmarketing » qui porte sur les différents courriels envoyés aux clients.

Le volume des courriels envoyés aux clients en « push » (en lien avec les options « Opt-in » et « Opt-out »¹⁵⁶) étant le plus conséquent, nous les avons traités en premier (contrairement à ce qui figurait sur l'ébauche de la boîte à outils, cf. annexes du chapitre 6). Nous avons obtenu des professionnels surtout des informations concernant comment les offres des courriels en « push » sont constituées. Le professionnel du marketing spécifie

d'abord différents profils de clients à partir de deux dimensions : les rayons dans lesquels ils achètent et leur niveau de fidélité ou d'activité d'achats (« *C'est le niveau de commandes du client, c'est si un client qui a tendance à dormir et qui commande plus ou...* »). Les offres sont ensuite créées à partir de ces profils d'acheteurs et à partir des événements de l'actualité (ex. sortie du dernier film « Machinchose »). Certains courriels, ainsi dits personnalisés, sont envoyés (en « push ») à ces profils de cyber-consommateurs et d'autres offres plus généralistes (ex. moins 20% sur le CD du très renommé groupe « Trucmuche » en tournée en France) sont envoyées à tous les clients de la base de données.

Concernant ces différents courriels en « push », nous avons proposé d'évaluer :

- la satisfaction des internautes concernant leur contenu en termes de produits,
- l'intérêt des informations les accompagnant,
- et la quantité globale de ces courriels.

Le dernier point présente un caractère majeur. En effet, si 20% seulement des internautes de Singapour (Teo, 2002) disent supprimer systématiquement tout courriel de « spam », ils sont plus de la moitié aux Etats-Unis et en Europe à le faire (Kehoe *et al.*, 1997). De plus, les internautes occidentaux sont une grande majorité à être fortement en désaccord avec toute opération de *mass mailing*, avec tout réseau publicitaire qui compile des informations de plusieurs sites marchands et avec l'envoi de courriels à partir de leurs données démographiques (ex. date d'anniversaire) (Kehoe *et al.*, 1997).

Un recueil par questionnaire libre (questions ouvertes) a également été proposé pour connaître les divers souhaits et suggestions des internautes, comme par exemple, une liste d'enregistrement des internautes qui désirent ne pas recevoir de courriels non-sollicités ou une liste des « spammeurs » connus pour permettre de filtrer l'envoi de leurs courriels (Kehoe *et al.*, 1997).

Les « Newsletters », type « lettres d'actualités », quant à elles, relèvent d'une démarche sans ambiguïté et volontaire de la part de l'internaute de recevoir de l'information ou des offres. Même si ces « Newsletters » ne sont a priori pas personnalisées, la cible des internautes les recevant est mieux définie et une enquête concernant cette cible peut avoir lieu. Le contenu de l'enquête demeure le même que précédemment (cf. tableau ci-dessus). Seule la question sur la quantité de courriels envoyés est supprimée, car le nombre de « Newsletters » est habituellement plus raisonnable.

Enfin, nous avons distingué les « Newsletters » personnalisées des précédentes « lettres d'actualités ». Pour celles-ci, l'internaute doit choisir des centres d'intérêt comme le jazz et le rock, en musique, la cuisine et les polars, en littérature, etc. Il recevra ainsi les dernières informations et offres du site marchand concernant ces rubriques de produits. La même évaluation que la précédente a été proposée (cf. tableau ci-dessus).

Les deux types de « Newsletters » ne sont pas développés par le Sitemarchand.com, mais étant donné que nous cherchons à généraliser notre outil d'aide à toute entreprise de commerce électronique, nous les avons ajoutés.

Enfin, nous avons précisé une rubrique « Aspects visuels de tout courriel » qui renvoie les professionnels à la catégorie « Look & co ».

Compréhension de la page « compte perso »

Tableau 72 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Compréhension de la page "compte perso" »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Compréhension de la page « compte perso »	<p>Les internautes ont-ils immédiatement identifié les « pages perso » (création de compte, services personnalisés) et à quoi elles servent ?</p> <p>Les internautes trouvent-ils les « pages perso » utiles et les offres personnalisées sur ces pages intéressantes ?</p>	<p>Il s'agit ici de la page « compte perso », « mon compte » ou encore « espace clients », « service clients », etc. qui rassemblent un ensemble de fonctionnalités ou services personnalisés (cf. liste ci-dessous).</p> <p>Appréciation (plutôt par Q) de :</p> <ul style="list-style-type: none"> • la facilité à identifier la page « compte perso » sur le site (depuis la page d'accueil ou autres), • la clarté de l'objectif de la page « compte perso » et des services associés (Avez-vous rapidement compris ce à quoi sert la page « compte perso » ? A quoi sert la création d'un compte ? A quoi sert le choix de centres d'intérêt ? etc.) - Compréhension des libellés de ces services et des explications succinctes les accompagnant, • l'utilité de la page ou compte « perso » (très utile, peu, etc.) et l'intérêt des offres personnalisées sur ces pages. <p>Cf. aussi catégories « Page d'accueil – rubriques et services principaux », « Logistique/Services - utilité des services » et « Faciliter à commander un produit – création d'un identifiant ou compte ».</p> <p>Liste des services personnalisés possibles :</p> <ul style="list-style-type: none"> • suivi de commande, • choix de centres d'intérêt pour un système de favoris internes (les rubriques choisies apparaissent sur les pages perso, ex. jazz et rock en musique, polar et cuisine en littérature, etc.), • présélection de produits en attente, etc.

Cet élément a été ajouté. Il n'y figurait pas dans l'ébauche de la boîte à outils (cf. annexes du chapitre 6). Pourtant les professionnels du Web marchand sont très concernés par cet élément et ont besoin de données sur l'appréciation subjective de cette page « compte perso ». Tout d'abord, le responsable du Webmastering nous explicite ce dont il s'agit : « C'est-à-dire que quand vous allez dans la page [compte perso], il y a des liens directs sur ses rubriques préférées, un lien direct sur le roman policier, un lien direct sur la science fiction... ». Cette page « compte perso » contient tous les autres services personnalisés du site : suivi de commande, accès à une présélection de produits en attente (pour les retrouver lors d'une prochaine connexion), etc.

Le professionnel du marketing et le responsable de la charte graphique sont soucieux de savoir :

- si l'on identifie facilement l'accès à cette page « compte perso » depuis toutes les pages du site,
- si l'internaute identifie rapidement à quoi sert cette page, la création d'un compte et les services associés (à partir de son intitulé, des explications très succinctes l'accompagnant ou des intitulés des services qu'elle contient),
- et enfin si cette page et la création d'un compte sont utiles et si les offres personnalisées sont intéressantes.

Cette évaluation peut être complétée par des évaluations de différentes catégories ; nous y avons renvoyé les professionnels (cf. tableau ci-dessus). Elle est parfois redondante avec ces catégories mais, une fois de plus, nous l'avons maintenu avec tous ses détails, étant donné son importance pour les professionnels du Web marchand.

De plus, tous les sites marchands ne sont pas constitués de la même manière, même s'ils ont une forte tendance à se standardiser. Quelques sites proposent encore le suivi de commande, la présélection de produits en attente, etc. sans passer par une « page perso » qui rassemble ces fonctionnalités ; d'où l'intérêt de créer aussi les rubriques « Page d'accueil – rubriques et services principaux », « Logistique/Services – utilité des services » et « Faciliter à commander un produit – création d'un identifiant ou "compte perso" » qui permettent alors de couvrir l'aide à l'évaluation pour tout site marchand.

3.2.8. Catégorie « Logistiques & Services »

3.2.8.1. Utilité des services

Tableau 73 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Utilité des services »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Utilité des services	Quels sont les services les plus utiles/utilisés par les internautes ?	Etude par Q. Liste de réponses possibles : <ul style="list-style-type: none"> • suivi des commandes (historique des commandes en cours), • création d'une présélection de produits à acheter, • création d'une liste de souhaits accessibles à des proches, • réception d'alertes sur les sorties de nouveaux produits, • enregistrement des coordonnées personnelles pour les conserver d'une commande à une autre, etc.

Cet élément a été créé à partir du questionnaire Noël 2002 et à partir des propos du responsable de la charte graphique qui nous rend attentif au fait qu'il faut interroger les internautes quant à l'utilité des différents services associés à l'achat (suivi des commandes, réception d'alertes sur les sorties de nouveaux produits, etc.).

L'élément « Utilité des services » complète l'élément « Rubriques et Services principaux » de la catégorie « Page d'accueil ». En effet, après avoir vérifié si les internautes identifient et comprennent bien les libellés des services, il s'agit de savoir lequel est le plus utile pour eux (pour le mettre en avant, pour l'améliorer sur le plan logistique, etc.).

Après cette évaluation globale, nous avons proposé des évaluations plus ou moins approfondies qui permettent de creuser divers aspects concernant chaque fonctionnalité ou service logistique.

3.2.8.2. Suivi de commande

Tableau 74 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Suivi de commande »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Suivi de commande	Les internautes ont-ils le sentiment de pouvoir bien suivre l'acheminement de leur commande ? Les annonces dans les messages leur semblent-elles fidèles à la réalité ?	Appréciation générale (plutôt par Q) de l'efficacité du suivi de commande (très efficace, efficace, etc.). EA (par Q ou par E), appréciation : <ul style="list-style-type: none"> • du nombre de courriels ponctuant le parcours du produit commandé, • éventuellement, de l'adéquation des annonces par rapport à la date d'arrivée effective du produit, dans les courriels ou dans les messages sur le panneau de suivi de commande. Facilité de visualisation du service (cf. catégorie « Page d'accueil – rubriques et services principaux ») et de son utilisation (cf. catégorie « Facilité à commander un produit »).

Avec le panier, la fonctionnalité de suivi de commande est la plus importante, aux yeux des professionnels du marketing et du Webmastering (« une fonction qui nous paraît très importante est le suivi de commande »). Pour un site marchand, le suivi de commande doit, en effet, être soigné (Liu & Arnett, 2000). Ainsi, le professionnel de la veille-marketing nous confirme l'intérêt de nos diverses propositions d'évaluations. Il s'agit, dans un premier temps, de connaître la satisfaction générale des internautes quant à l'efficacité du suivi de commande. Si les résultats sont mauvais, une évaluation approfondie peut ensuite vérifier si les courriels envoyés aux clients ne sont pas assez nombreux et rapprochés les uns des autres ou si le contenu des annonces n'est pas en adéquation avec la date d'arrivée du produit.

Si ces évaluations semblent simples, il faut savoir que leurs résultats peuvent engendrer des modifications plus ou moins considérables. Trois pistes de modification sont à envisager :

- augmenter simplement le nombre de courriels qui ponctuent les différentes étapes du parcours du produit (ce qui suppose un reparamétrage du logiciel d'envoi de courriels, voire de réviser le contrat avec l'éventuel sous-traitant réalisant l'envoi de ces courriels) et/ou mieux estimer le temps moyen d'arrivée du produit pour faire une annonce plus adéquate,
- modifier les modules de gestion utilisés pour donner une visibilité au cyber-consommateur de l'avancement de sa commande, c'est-à-dire l'accès sur le site à un panneau de suivi ou la réception de courriels annonçant le traitement de la commande, le départ du produit, etc.
- ou enfin optimiser la chaîne de distribution elle-même de façon à ce que les produits partent effectivement à la date annoncée (cf. « Garanties de livraison », dans le paragraphe suivant).

3.2.8.3. Garanties de livraison

Tableau 75 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Garanties de livraison »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Garanties de livraison	Les internautes sont-ils satisfaits de la date/moment d'arrivée de leur produit au regard du délai annoncé ? Eventuellement estiment-ils que le produit soit arrivé en bon état ?	Selon le choix de transporteurs des sites : La Poste, transporteur interne/externe. Eventuellement, prévoir différents segments de clients : internautes à l'étranger ou non. Appréciation (plutôt par Q) : <ul style="list-style-type: none"> • de l'arrivée du produit par rapport au délai annoncé (plus tôt que prévu, conformément aux délais annoncés, etc.), • éventuellement de l'état du produit à l'arrivée. Corrélation avec les remontées clientèles : nombre de réclamations déposées concernant les retards, les produits défectueux (engendrant un retour et un échange de produits).

Le questionnaire Noël 2001 et les propos du professionnel de la veille-marketing nous assurent que cet élément constitue un besoin en évaluation, même quand un commerce électronique ne maîtrise pas l'expédition des colis de bout en bout (expédition via La Poste).

Il s'agit là d'estimer si les internautes sont satisfaits des délais de livraison et éventuellement du bon état du produit à son arrivée. Il arrive que certains sites marchands fassent des enquêtes à ce sujet (y compris s'ils traitent avec La Poste) et publient les résultats sur leur site (ex. pourcentage d'internautes ayant été livrés dans les délais annoncés).

Concrètement, dans les délais de livraison, sont cumulés les délais de disponibilité du produit (aucun si le produit existe dans les stocks des entrepôts de l'entreprise), les délais de mise en colis (ex. dépendant de l'existence d'une ou plusieurs machines de colisage) et enfin des délais de livraison (dépendant du transporteur interne ou externe à l'entreprise et du lieu/pays d'arrivée du produit). Même si tous ces délais ne sont pas maîtrisables par l'entreprise marchande (ex. concernant La Poste), l'objectif d'une évaluation peut, par exemple, être de vérifier si la satisfaction des internautes s'améliore suite à l'investissement d'une machine de colisage supplémentaire. Il peut aussi s'agir d'interroger des clients étrangers si le site marchand fait de la vente à l'international. Ces enquêtes sont généralement décidées après qu'un nombre jugé important de réclamations soit parvenu au service clients (remontées clientèles).

Une fois de plus, la volonté d'améliorer les délais d'expédition conduit à considérer différents éléments en fonction des caractéristiques de l'entreprise marchande. Divers dispositifs existent pour faire face à ce « casse-tête » que représente la livraison, en raison de l'exigence grandissante des internautes (Macarez & Leslé, 2001). Ainsi, si l'entreprise marchande a ses propres ressources d'expédition (transporteur interne), il faut revoir l'organisation de travail ou les divers investissements (ex. machines de colisage) de la chaîne de distribution, depuis les entrepôts jusqu'à l'arrivée du produit « dans le panier de la ménagère ». Sinon, depuis les entrepôts jusqu'à la levée du produit vers La Poste ou tout autre transporteur externe. Dans ce dernier cas, l'entreprise marchande ne peut pas maîtriser la performance de ces sous-traitants qui prennent le relais, si ce n'est d'en changer. Enfin, certaines entreprises marchandes confient toute la chaîne de distribution depuis le fabricant ou fournisseur jusqu'au panier de la ménagère à des « e-logisticiens » (Macarez & Leslé, 2001), s'affranchissant ainsi de gérer eux-mêmes des entrepôts et tout équipement de colisage, d'étiquetage, etc. Dans ce cas, l'entreprise marchande peut aussi mener une telle enquête auprès de ses clients et ensuite exiger de ses sous-traitants d'analyser les points faibles de leur prestation, s'il y en a.

Augmenter la réactivité de tout ce système peut aussi supposer la révision des progiciels de l'entreprise marchande de gestion des commandes, de planification des approvisionnements auprès des différents fournisseurs, etc. (Macarez & Leslé, 2001).

3.2.8.4. Possibilité de RDV & Relais colis

Tableau 76 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Possibilité de RDV & Relais colis »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Possibilité de RDV & Relais colis	<p>Les internautes sont-ils satisfaits des services de livraison (par rendez-vous, par relais colis) ?</p> <p>Est-ce réellement utile ?</p> <p>La plage horaire de RDV est-elle adéquate ?</p> <p>Les horaires d'ouverture des relais et leur proximité sont-ils satisfaisants ?</p>	<p>Selon les sites, appréciation générale (plutôt par Q) :</p> <ul style="list-style-type: none"> • de la qualité du service de livraison (par rendez-vous, par relais colis), • éventuellement, de l'utilité de ces services par rapport à une livraison classique (ex. d'intérêt : retour du produit directement pris en charge par le relais colis sans frais, etc.). • Plus précisément, possibilité d'évaluer l'appréciation : <ul style="list-style-type: none"> ○ de l'arrivée du produit par rapport au RDV prévu, ○ ou de la satisfaction quant aux heures d'ouverture des relais colis, à la proximité des relais colis, etc. <p>Selon les résultats et en fonction de leurs capacités, les sites marchands seront susceptibles de prendre certaines décisions de réaménagement de leur infrastructure logistique, de révision des contrats avec les sous-traitants, etc.</p>

Pour tout dire, notre propre expérience d'achat par Internet avec des sites proposant des services de livraison particuliers comme la livraison à domicile avec rendez-vous ou un système de relais colis (qui sont des magasins-relais de proximité), nous a fait conserver cette rubrique, peu considérée par les professionnels du Sitemarchand.com. A l'époque de nos entretiens, l'entreprise marchande n'avait pas mis en place de tels moyens. Depuis, elle a développé la livraison par relais colis.

Diverses considérations soutiennent l'utilité de cette évaluation. La possibilité de livraison à domicile avec prise de rendez-vous ou l'existence de relais colis peut influencer sur la perception de la qualité de service, puisque les cyber-consommateurs sont des personnes particulièrement pressées (Bellman *et al.*, 1999 ; Lohse *et al.*, 2000). De plus, les relais colis permettent généralement de voir le produit, de vérifier ses dimensions, ses couleurs, etc. et de le retourner directement (sans aucun frais) s'il ne convient pas. Le professionnel de la veille-marketing l'atteste : « *Oui et c'est pour ça que les relais, ça marche mieux, parce qu'il suffit de reporter le paquet au relais. Ca c'est un des points intéressants...* ».

Nous avons proposé d'évaluer :

- la perception de la qualité globale de ces services pour s'assurer qu'ils rencontrent l'approbation des clients,
- éventuellement d'évaluer la perception de leur utilité (ou valeur ajoutée du point de vue du client) par rapport à une livraison classique,

- la satisfaction des internautes quant au moment d'arrivée du produit au regard du rendez-vous prévu,
- la satisfaction des internautes quant aux heures d'ouverture des relais colis ou encore quant à leur proximité (par exemple, il est assez désagréable de trouver la porte de la « boutique relais » fermée le Mercredi sans en avoir été prévenu depuis le site Internet).

Ecore une fois, ces dernières évaluations sont assez approfondies, elles ne seront mises en œuvre que si une vaste remise en question est envisagée. En effet, d'après le professionnel de la veille-marketing, une étude commerciale préalable est réalisée par l'entreprise marchande pour choisir une entreprise qui développe ce réseau de relais colis. Ainsi, une fois le sous-traitant choisi et le réseau de magasins de proximité établi, il n'est sans doute pas évident pour l'entreprise marchande de modifier tout cela.

3.2.8.5. Emballage cadeau

Tableau 77 : Outil d'aide à la conception d'évaluations de sites Internet marchands pour les professionnels du marketing et du Webmastering, concernant le besoin en évaluation « Emballage cadeau »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Emballage cadeau	Les internautes trouvent-ils utile de pouvoir bénéficier d'un emballage cadeau ?	Pour les sites proposant ce service (gratuit, la plupart du temps) : appréciation de l'utilité de l'emballage cadeau (très utile, utile, etc.)

Le questionnaire Noël 2001 et l'approbation du professionnel de la veille-marketing sont les raisons d'existence de cette rubrique : il s'agit d'évaluer si les internautes jugent utile l'emballage cadeau. Cette évaluation concerne les sites qui réalisent de tels emballages et veulent vérifier son intérêt.

3.2.8.6. Frais d'envoi

Tableau 78 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Frais d'envoi »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Frais d'envoi	Les internautes estiment-ils les frais d'envoi raisonnables ?	<p>Etude par Q. Eventuellement, prévoir différents segments de clients (internautes à l'étranger ou non car les frais d'envoi varient selon le pays d'expédition, voire le département d'expédition).</p> <p>Appréciation générale de la perception des frais d'envoi (très chers, chers, bon marché, etc.).</p> <p>Eventuellement en comparaison avec la concurrence, cf. catégorie « Offre – Positionnement ».</p>
Somme minimum d'achat pour une livraison gratuite	Quelle est leur estimation de la somme à ne pas dépasser pour les frais d'envoi ou la somme minimum d'achat pour une livraison gratuite ?	<p>EA par FG (les FG sont particulièrement indiqués pour la recherche d'un consensus autour d'un « bon » prix) auprès d'internautes ayant réalisé des comparaisons de plusieurs sites marchands d'un même secteur de vente (ou les réalisant à l'occasion de l'évaluation).</p> <ul style="list-style-type: none"> • Estimation des frais d'envoi à ne pas dépasser (ex. recherche de consensus, en groupe, autour d'un prix), • Estimation de la somme à ne pas dépasser concernant la somme minimum d'achat pour une livraison gratuite. <p>Cette EA pour tout commerce traitant avec La Poste ou des transporteurs internes/externes, supposant éventuellement une remise en question assez avancée concernant des choix logistiques ou commerciaux (ex. amélioration de la rentabilité du poste de colisage, etc.).</p>

Le professionnel de la veille-marketing aurait vu rassembler les prix des produits et les frais d'envoi (« *La question du prix, c'est le prix du produit et le prix du port* »). Cependant, étant donné que les frais d'envoi sont dépendants de la somme totale d'achat, du pays, voire du département d'expédition, nous avons suggéré de traiter ce point à part.

Pour estimer la satisfaction des internautes quant aux frais d'envois, nous avons proposé une évaluation comparative des frais pratiqués par plusieurs sites marchands. En effet, la comparaison de quatre sites marchands d'un même secteur de vente montre que pour le même type de produit et un même service postal, le tarif n'est pas le même¹⁵⁷. De même, la plupart des sites pratiquent des politiques de suppression des frais de livraison à partir d'une certaine somme d'achat qui varie de l'un à l'autre.

Après une évaluation généraliste, si les résultats s'avèrent mauvais, nous avons proposé de procéder à une évaluation (marketing) plus approfondie pour connaître quelle est la somme à ne pas dépasser pour les frais d'envoi ou même pour la somme minimum d'achat pour une

livraison gratuite. Les frais d'expédition doivent, en effet, rester convenables, car les internautes ont une représentation de ce qui est raisonnable et trouvent certains frais inacceptables (Nielsen *et al.*, 2001).

Nous avons proposé une évaluation par *focus groups* qui se prêtent très bien à la recherche d'un consensus (cf. chapitre 3) d'un « bon » prix. En effet, les discussions de confrontation et les négociations dans un groupe obligent les participants à se projeter dans une situation plus proche du réel, c'est-à-dire à réfléchir au prix que réellement ils ne voudraient pas dépasser. Dans les entretiens individuels, en revanche, la réponse à ce type de question est rarement réaliste.

3.2.8.7. Retour de produits

Tableau 79 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Retour de produits »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Retour de produits	Les internautes estiment-ils les modalités de retour de produits acceptables ?	<p>EA plutôt par Q (susceptible d'engendrer une remise en question assez importante au niveau de l'infrastructure logistique).</p> <p>Appréciation du sentiment général quant aux modalités de retour de produits (simples ; compliquées mais acceptables ; pénibles ; insupportables, je ne commande pas par Internet à cause de cela, etc.).</p> <p>Il s'agit de vérifier que cela ne soit pas un frein rédhibitoire à consommer sur le site Internet.</p> <p>Liste des clauses de retour :</p> <ul style="list-style-type: none"> • en cas de changement d'avis du client dans le cadre de la clause « satisfait ou remboursé » : frais de renvoi habituellement à la charge du client, retour du produit avec son emballage intact et sa facture, renvoi par La Poste ou par prise de RDV avec le transporteur, etc. • ou en cas d'arrivée du produit en mauvais état ou défectueux : idem mais frais de renvoi remboursés.

Les modalités de retour de produits très contestées par les internautes sont pourtant comparables à celles généralement pratiquées par les « VPCéistes »¹⁵⁸ traditionnels. Les acheteurs habitués à la vente par correspondance sont plus familiarisés avec ces procédures (retour du produit intact avec son emballage et la facture, le tout renvoyé aux frais de l'acheteur, sauf en cas d'erreur du commerçant). Ainsi, l'émergence du commerce électronique a pu attirer une clientèle moins accoutumée aux contraintes de l'achat à distance¹⁵⁹ et donc plus exigeante. Deuxièmement, une étude par *focus groups* concernant les freins et facteurs d'achat des cyber-consommateurs (réalisée pour la construction de l'outil

M.A.S.C. décrit au chapitre précédent ; non publiée) a montré que l'acte d'achat est perçu comme plus rapide sur Internet que par V.P.C. Un autre facteur, plus inconscient, peut être relatif au monde virtuel et à la magie de faire venir à soi les objets qu'on désire : commander les produits depuis chez soi, à n'importe quel moment du jour et de la nuit sans se déplacer. Le fait de devoir repositionner le produit et ses éventuels accessoires dans l'emballage, de « re-fixer » le paquet, de le porter et de patienter au bureau de poste contraste fortement avec tout imaginaire de magie. La facilité des procédures de retour et d'échange des produits est, en effet, assez capitale pour les internautes (Tilson *et al.*, 1998).

Ainsi, un questionnement ouvert sur ce point lors d'une évaluation donnerait sans doute lieu au recueil d'un foisonnement d'épanchements divers que nous avons déjà entendus lors de nos entretiens utilisateurs. En fait, nous doutions de l'utilité de cette évaluation puisque, si les commerçants offrent à leurs clients la possibilité du fameux délai « satisfait ou remboursé »¹⁶⁰, ils ne peuvent pas, dans la plupart des cas, prendre à leurs charges (déplacements, frais d'envoi) les changements intempestifs d'avis des clients. Et si les produits parviennent défectueux aux clients, les frais de renvoi leur sont alors remboursés et le produit échangé. Cependant, le professionnel de la veille-marketing nous maintient que : « *Ben si ! Il faut quand même valider que la procédure fonctionne bien correctement aux yeux des clients* ». En effet, quelques internautes de nos entretiens utilisateurs disent ne plus acheter sur le Web en raison de ces éventuelles complications.

Nous avons suggéré d'évaluer à quel point c'est un frein : « *Est-ce que c'est un frein, est-ce que c'est devenu un frein ? Est-ce que ça évolue dans le bon sens ?* » Le questionnement proposé (cf. tableau ci-dessus) cherche à interroger les internautes en ce sens.

3.2.8.8. Service de maintenance

Tableau 80 : Outil d'aide à la conception d'évaluations de sites Internet concernant le besoin en évaluation « Service de maintenance »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Service de maintenance	Les internautes sont-ils satisfaits du service de maintenance ?	EA. Pour les sites disposant d'un centre interne de service après-vente réalisant des opérations de maintenance des produits vendus ou éventuellement faisant appel à des sous-traitants en la matière. Appréciation générale (plutôt par Q) de l'efficacité du service de maintenance. Éventuellement, appréciation de : <ul style="list-style-type: none"> • l'efficacité des opérations de maintenance (réparation ou remplacement du produit), • éventuellement, la convivialité et l'accueil des opérateurs.

Le professionnel de la veille-marketing nous confirme l'utilité d'évaluer la satisfaction des clients concernant les prestations du service de maintenance (l'efficacité du service et la convivialité des opérateurs) : « *Ben, ils ont des accords avec des installateurs, il y a un service après vente... de maintenance. D'accord, mais donc, ils iraient évaluer la satisfaction d'un client pour le répercuter sur ce qui est sous-traité ? Tout à fait, oui.* ». Ainsi, cette évaluation concerne les entreprises marchandes ayant leur propre service de maintenance interne, mais aussi celles qui le sous-traitent.

Les entreprises marchandes peuvent alors être amenées à revoir leur choix de prestataires ou peuvent les encourager à analyser et améliorer leurs points faibles, de même que pour leur sous-traitant de livraison.

3.2.9. Catégorie « Interactivité »

Cette catégorie ne comprend plus que la question du contact directe avec un service clients ou tout autre entité de l'entreprise comme il a déjà été précisé (cf. catégorie « Netmarketing »).

3.2.9.1. Possibilité de contacter un service clients

Tableau 81 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Possibilité de contacter un service clients »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Possibilité de contacter un service clients	Les internautes (visiteurs, clients, etc.) ont-ils connaissance de pouvoir contacter un service clients ? Pour quel type de questions ou de réclamations ?	<p>EA (préalable) par E.</p> <p>Questionnement sur la connaissance des internautes déjà clients par rapport (en fonction des services assurés par le site) :</p> <ul style="list-style-type: none"> • au seul fait de pouvoir contacter un service clients (vérifier s'ils ont remarqué que c'est un service proposé par le site, s'ils ont le réflexe d'y penser – ce qui n'est a priori pas toujours le cas), • aux raisons pour lesquelles ils peuvent appeler le service clients (ex. citation libre des différentes possibilités connues, cf. liste ci-dessous). <p>Corrélation avec des données d'analyse des remontées clientèles (pour quelles raisons majeures les internautes appellent-ils le service clients ?).</p> <p>Liste des services possibles :</p> <ul style="list-style-type: none"> • assistance technique, • achat par téléphone, • réclamation par rapport au paiement, • réclamation par rapport à un retard de livraison, • retour de produits, • conseil pour un achat, etc. <p>Rq : Visibilité des informations sur les contacts, cf. catégorie « Look & co » et Compréhension des informations, cf. catégorie « Informations associées au service de vente ».</p>

Après réflexion, nous avons maintenu cet élément (en tant qu'évaluation approfondie), même s'il ne fait pas l'unanimité. En effet, nous avons été surpris par le nombre d'internautes de nos entretiens utilisateurs, plus novices comme plus experts, n'ayant apparemment pas le réflexe d'appeler un service clients, simplement du fait qu'ils utilisent Internet. Cela rejoint sans doute l'imaginaire du monde virtuel auquel nous avons déjà fait référence.

En outre, plus de la moitié des internautes préfèrent ou requièrent un site marchand qui ait un magasin ou une quelconque présence « terrestre » permettant de commander par courrier comme avec les « VPCéistes » classiques (Kehoe *et al.*, 1997). Cela est peut-être dû au fait que les cyber-consommateurs font davantage confiance aux entreprises marchandes qu'ils connaissent déjà (Jarvenpaa & Todd, 1997; Jarvenpaa *et al.*, 2000). Ainsi, contacter un service clients, surtout quand le site est un « pure player », peut contribuer à rendre la relation au site marchand plus réaliste et à rassurer l'internaute.

Nous avons proposé de vérifier si les clients du site ont ce réflexe là ou non et, dans un second temps, si nécessaire, de rechercher comment améliorer ce réflexe. Nous avons proposé d'estimer le nombre d'internautes non pas seulement ayant vu le lien du service clients sur le site, mais ayant aussi le réflexe d'appeler le service clients, si nécessaire. On peut, de même, les interroger sur leurs connaissances concernant les raisons pour lesquelles ils peuvent appeler le service clients.

Ces données peuvent, une fois de plus, être corrélées à des données objectives concernant la quantité d'appels pour chaque cause (problème technique relatif au site, réclamation par rapport au paiement, etc.). Des outils permettent de classer les messages par thème et de constituer une base de données (Macarez & Leslé, 2001).

Enfin, une indication renvoie les professionnels aux catégories « Look & co » et « Informations associées au service de vente » concernant les aspects relatifs à la visibilité et à la compréhension de ces informations (cf. tableau ci-dessus).

3.2.9.2. Qualité du service clients

Tableau 82 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant les besoins en évaluation « Qualité du service clients »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Moyens pour contacter le service	<p>Les internautes sont-ils satisfaits de la facilité à joindre le service clients ?</p> <p>Ont-ils le sentiment d'accéder facilement à une adresse électronique ou à un n° de tél., à tout moment sur le site ?</p> <p>Faut-il d'autres moyens supplémentaires ?</p>	<p>En fonction des moyens mis à disposition par le cyber-commerce. Appréciation (plutôt par Q) :</p> <ul style="list-style-type: none"> • de la facilité à joindre le service clients, • de la facilité à trouver les coordonnées électroniques et téléphoniques, à tout moment sur le site. Cf. aussi catégorie « Look & co ». <p>EA par Q ou par FG (si volonté d'amélioration des services existants) sur les autres moyens à mettre en place pour satisfaire les internautes (ex. visiophonie).</p> <p>Liste des moyens (dans l'ordre du plus au moins courant) :</p> <ul style="list-style-type: none"> • messagerie électronique, • téléphone, • téléphone par Internet (« push-to-talk »), • visiophonie, etc.
Qualité des réponses apportées & Convivialité	<p>Les internautes sont-ils satisfaits des réponses apportées ?</p> <p>Eventuellement de la convivialité des télé-assistants ?</p>	<p>Appréciation (plutôt par Q) :</p> <ul style="list-style-type: none"> • de l'efficacité dans le traitement de leur requête, c'est-à-dire : <ul style="list-style-type: none"> ○ de la rapidité de la réponse (ex. nombre de jours avant réception du courriel, attente téléphonique avant d'obtenir un télé-assistant, télé-conseiller, etc.), ○ de la pertinence de la réponse (la réponse que vous avez obtenue a-t-elle résolu votre problème ?) • de la convivialité, de l'accueil des télé-assistants ou autres opérateurs.

Plus classiquement, nous avons proposé de chercher à savoir si les internautes sont satisfaits des dispositifs (facilité des moyens et facilité à les identifier) mis à leur disposition pour contacter le service clients, ainsi que de la rapidité et de la pertinence de la réponse et enfin de la convivialité des opérateurs. L'importance de ces éléments est confirmée dans la littérature (Liu & Arnett, 2000), à partir du questionnaire Noël 2001 et approuvée par les professionnels du (et de la veille-) marketing. Cependant, le professionnel du marketing nous explique que ce n'est pas une évaluation réalisée de façon permanente. Elle a été accomplie pour Noël 2001 mais pas pour Noël 2002, sans doute parce que le résultat était satisfaisant après la première évaluation.

L'évaluation approfondie proposée dans « Moyens pour contacter le service » concernant l'intérêt de la mise en place de dispositifs plus sophistiqués pour contacter le service clients

est assez particulière. Nous l'avons indiqué à titre indicatif pour les sites qui seraient concernés par cette problématique.

3.2.10. Catégorie « Sécurité & Paiement »

Nous avons déplacé cette catégorie « Sécurité – Paiement » après les catégories « Logistique » et « Interactivité », étant donné que les professionnels, en particulier du marketing, sont en réalité davantage préoccupés de la logistique et des possibilités d'interactivité entre les clients et le commerce électronique, que de la sécurité.

3.2.10.1. Facilités de paiement

Tableau 83 : Outil d'aide à la conception d'évaluations de sites Internet concernant le besoin en évaluation « Facilités de paiement »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Facilités de paiement	Les internautes sont-ils satisfaits des différents moyens de paiement mis à leur disposition ?	Etude plutôt par Q. Appréciation générale : <ul style="list-style-type: none"> • de la satisfaction des internautes concernant les moyens de paiement proposés, • du moyen de paiement le plus apprécié (CB, chèque, mandat, fax, éventuellement porte-monnaie virtuel, solution S.E.T. cf. aussi rubrique « Options de sécurisation »). • Recueil éventuel de leurs attentes non réalisées (quels moyens sont manquants, insuffisants ? Ex. chèque à partir de 140€seulement).

Les moyens de paiement sont divers : différentes cartes bancaires (Mastercard, Visa, American Express, etc.), chèque à partir d'un montant minimum (ex. 0 € ou 140 €), mandat, communication du numéro de carte bancaire ou de cartes privées par téléphone, par fax, etc. Si la majorité des cartes bancaires est acceptée par tous les cyber-commerces, concernant les autres modalités, on observe des variations de l'un à l'autre. Or, il est recommandé de fournir la possibilité de paiement par chèque, de paiement à réception du produit, voire la solution du paiement étalé dans le temps ; cela, afin de réduire les inquiétudes concernant le paiement sécurisé et la peur de payer pour un produit qui n'est pas livré ou parvient détérioré (Nielsen *et al.*, 2001; Teo, 2002).

Nous avons proposé de vérifier le niveau de satisfaction des internautes concernant les moyens de paiement proposés, de connaître le moyen qu'ils apprécient le plus et

éventuellement de savoir s'ils ont certains souhaits et critiques à formuler. Le professionnel de la veille-marketing approuve cela.

3.2.10.2. Sentiment de sécurité

Tableau 84 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Sentiment de sécurité »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Sentiment de sécurité	Les internautes, au moment de la transaction, se sentent-ils en sécurité ?	Complémentaire à l'évaluation de l'effet des indices de sécurité de la page d'accueil (catégorie « Page d'accueil - Réassurance »). Etude plutôt par Q : <ul style="list-style-type: none"> • Appréciation du sentiment de sécurité de l'internaute, au moment de la transaction (Vous êtes vous senti en sécurité au moment de passer la commande ? Aviez-vous quelques incertitudes quant à la sécurité de la transaction sur ce site ?). • Questionnement sur les différents éléments (autres que ceux de la page d'accueil) qui ont participé à ce sentiment de sécurité (ex. lien depuis le panier vers les pages d'explications sur la sécurité).

Le questionnaire Noël 2001 interroge les internautes sur leur sentiment de sécurité au moment de passer la commande et les raisons pour lesquelles ils se sont sentis en confiance. Ce questionnement n'est pas réitéré dans le questionnaire Noël 2002, sans doute, comme le précise le responsable du Webmastering, car il semble que la sécurité ne soit plus la préoccupation première des cyber-consommateurs du Sitemarchand.com. Ce déclin est confirmé par quelques travaux (Bellman *et al.*, 1999; Kehoe *et al.*, 1998b).

Nous avons néanmoins suggéré d'interroger le sentiment général de sécurité de l'internaute, au moment de la transaction. Cet élément complète la rubrique « Page d'accueil – Réassurance ».

De plus, nous avons conseillé un questionnement supplémentaire cherchant à savoir quels éléments ont participé à cela, de même que le suggère le professionnel du marketing au moment de traiter la rubrique « Réassurance ». Il ne s'agit plus de recueillir des impressions concernant les éléments de la page d'accueil, mais concernant tout autre élément de l'interface comme, par exemple, un lien depuis le panier vers les pages d'explications sur la sécurité.

3.2.10.3. Confidentialité

Tableau 85 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Confidentialité »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Confidentialité	<p>Quel est le sentiment des internautes quant au « niveau » de confidentialité assuré pour toutes leurs informations personnelles ?</p> <p>Ces engagements augmentent-ils le sentiment de sécurité ?</p>	<p>Appréciation générale par Q :</p> <ul style="list-style-type: none"> • de leur sentiment (impression de transparence) concernant la confidentialité assurée par le site, • et pour savoir si leur sentiment de sécurité augmente du fait de ces engagements. <p>Liste des engagements possibles de confidentialité :</p> <ul style="list-style-type: none"> • non-divulcation à un tiers des informations personnelles collectées (parcours sur le site, achats effectués), • enregistrement des fichiers-clients de l'entreprise à la C.N.I.L. (normalement obligatoire), • possibilité pour l'internaute de modifier les informations personnelles qu'il a communiquées (adresse, téléphone., etc.) ou de les supprimer.

Le responsable du Webmastering maintient que les internautes sont dorénavant plus attentifs à leurs données personnelles : « Là où ils font plus attention, c'est tout ce qui est relatif à la C.N.I.L.¹⁶¹ et à la préservation de leurs coordonnées perso... ». Une étude révèle, en effet, que la majorité des internautes (a) préfèrent avoir accès à leurs informations personnelles, (b) exigent que les sites fournissent une politique explicite concernant les données personnelles recueillies, (c) sont vivement opposés à ce que des informations puissent être revendues et enfin (d) souhaitent de nouvelles lois sur les questions de confidentialité sur le Web (Kehoe *et al.*, 1997). Ces inquiétudes ne sont sans doute pas que le fruit d'une phobie imaginaire, puisque 45% seulement des grands sites d'e-business français ont déclaré leurs fichiers à la C.N.I.L. alors que c'est normalement obligatoire (Macarez & Leslé, 2001).

Le professionnel de la veille-marketing nous confirme l'intérêt de l'évaluation proposée, à savoir évaluer le sentiment des internautes quant à la confidentialité des données personnelles, et aussi observer l'évolution de leur sentiment de sécurité (est-ce qu'il augmente du fait de l'existence de ces engagements ?). Ce professionnel nous précise que cette évaluation peut se réaliser par simple questionnaire et qu'il n'est pas nécessaire de trop l'approfondir : « Je ne sais pas si c'est nécessaire de rentrer dans une étude plus approfondie quand même. Ben disons que ça, ça peut être inclus dans l'enquête de...[...] Il n'empêche qu'une enquête suffit, quoi. Parce qu'il n'y a pas besoin, enfin, c'est assez simple comme question... comme questionnement ».

3.2.10.4. Options de sécurisation

Tableau 86 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Options de sécurisation »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Options de sécurisation	<p>L'option de non-archivage du numéro de carte bancaire ou la transaction sécurisée par une banque augmentent-ils le sentiment de sécurité des internautes ?</p> <p>D'autres techniques de paiement plus « pointues » augmentent-elles le sentiment de sécurité des internautes ?</p>	<ul style="list-style-type: none"> • Non-archivage du numéro de CB: si le numéro de CB n'est pas archivé, l'internaute doit le rentrer à chaque transaction ; si le numéro est archivé, l'internaute peut faire un achat rapide (enregistrement de toutes ses coordonnées) mais son numéro de CB est conservé sur les fichiers de l'entreprise qui présentent une certaine vulnérabilité (possibilité de pénétration de l'environnement informatique de l'entreprise à partir de son site Internet), • Transaction sécurisée par une banque : si le site est associé à une banque qui prend intégralement en charge la transaction, le numéro de carte est crypté mais en plus il n'est communiqué qu'à la banque elle-même qui notifie (ou non) son accord pour la transaction au commerçant. <p>Appréciation par Q :</p> <ul style="list-style-type: none"> • de leur connaissance concernant les précédents éléments, • de leur sentiment de sécurité (augmente-t-il du fait de l'existence des précédents éléments ?). <p>Etude de fond possible, à titre d'informations générales, concernant d'autres techniques de paiement plus « pointues » :</p> <ul style="list-style-type: none"> • S.E.T. (Secure Electronic Transaction) ou paiement avec la signature électronique de l'acheteur (clef cryptée fournie par la banque), donc authentification des coordonnées bancaires et du propriétaire réel de la carte directement par la banque qui notifie (ou non) son accord au commerçant, • Porte-monnaie virtuel, • Carte à puce. <p>Etude approfondie par FG : mise en évidence de tous les aspects qui contribuent au sentiment de sécurité (partenaires financiers ; solutions de paiement, de sécurisation et de confidentialité des informations personnelles ; slogans et informations formulés, etc.).</p>

Cette évaluation est très spécifique ; elle concerne un site marchand qui propose de ne pas archiver le numéro de carte bancaire pour plus de sécurité (cf. explication dans le tableau ci-dessus) et veut vérifier si l'internaute est ainsi rassuré. Il s'agit donc de vérifier si les internautes sont au courant de cette option de non-archivage et simultanément d'interroger leur sentiment de sécurité. Le professionnel de la veille-marketing valide cela.

Cette évaluation peut être étendue à tout site qui s'est associé à une banque qui prend intégralement en charge la transaction. Cela signifie que l'internaute saisit son numéro de carte bancaire via le serveur sécurisé de la banque et ce numéro n'est communiqué qu'à la

banque elle-même qui notifie (ou non) son accord pour la transaction au commerçant. Cela évite que le numéro de carte bancaire soit traité par le commerçant lui-même (ex. envoi de ce numéro par fax non sécurisé à la banque pour procéder à la transaction). Il s'agit, de même, d'interroger le sentiment de sécurité des cyber-consommateurs en lien avec cette prise en charge de la transaction par une banque.

De plus, étant donné que les évolutions en matière de moyens de paiement et de sécurisation sur le Web sont en perpétuelles évolutions, une évaluation plus généraliste peut (ou pourra) être envisagée concernant d'autres moyens de paiement/sécurisation plus « pointus », mais moins développés à l'heure actuelle. Il s'agit de la solution S.E.T. (*Secure Electronic Transaction* - cf. explication dans le tableau ci-dessus), du porte monnaie virtuel et de la carte à puce (Fouchard, 2001; Ghernaoui-Hélie, 2002; Macarez & Leslé, 2001; Teo, 2002). Cette étude, nécessitant probablement une co-réalisation par plusieurs entreprises, peut apporter des connaissances pour savoir si c'est utile de promouvoir de tels moyens de paiement et comment.

Enfin, une évaluation par *focus groups* peut chercher à mettre en évidence ce qui contribue le plus à rassurer l'internaute pour passer une transaction commerciale : les partenaires financiers ; les solution(s) de paiement, de sécurisation et de confidentialité des informations personnelles ; les slogans et informations formulés, etc. Cette évaluation qualitative doit permettre de donner des recommandations d'amélioration adaptées aux caractéristiques singulières du site marchand.

3.2.10.5. Partenaires financiers et organismes de contrôle

Tableau 87 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Partenaires financiers & Organismes de contrôle »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Partenaires financiers & Organismes de contrôle	Les internautes connaissent-ils les partenaires financiers et organismes de contrôle du site marchand ?	EA par Q. Appréciation de : <ul style="list-style-type: none"> leur connaissance des éventuels partenaires financiers (« tiers de confiance » assurant la transaction) ou organismes de contrôle d'un site (étude type évaluation de la notoriété d'une marque, en spontané ou en assisté), leur satisfaction concernant le choix des partenaires et l'efficacité du contrôle (ex. vérification des engagements de confidentialité) ou l'efficacité des assurances proposées (ex. remboursement en cas de fraude), pour savoir si leur sentiment de sécurité augmente.
	En apprécient-ils le choix et trouvent-ils leur action efficace ?	
	Ces divers partenaires augmentent-ils leur sentiment de sécurité ?	EA par E avec utilisation et lecture des pages explicatives de plusieurs sites marchands pour une évaluation comparative qualitative (plus riche).

Cet élément est dans la continuité de ce qui précède. Le professionnel de la veille-marketing l'approuve également. Nous avons proposé de vérifier si les clients ont connaissance des partenaires financiers (ex. banque prenant en charge le remboursement en cas de détournement du numéro de carte bancaire) et des labels et organismes de contrôle (intervenant en cas de non-remboursement après retour de produits défectueux, vérifiant le respect des engagements de confidentialité du site marchand, etc.) auxquels s'est associée l'entreprise marchande.

Après cette vérification, il est possible de savoir à quel point les clients sont satisfaits du choix de ces partenaires et de leurs engagements et prestations.

Enfin, il s'agit de vérifier si l'existence de ces derniers participe au sentiment de sécurité de l'internaute. Les sites audités par un organisme officiel seraient, en effet, préférés par les internautes (Kehoe *et al.*, 1997).

Les internautes n'ayant pas forcément une connaissance détaillée de toutes les actions des partenaires financiers et des organismes de contrôle, nous avons également proposé une évaluation par entretiens avec utilisation de plusieurs sites marchands et lecture comparée de leurs pages d'explications. Ce type d'évaluation qualitative et comparative est, en effet, généralement riche en données précises sur les préférences, impressions, inquiétudes, etc. des utilisateurs.

Concernant la rubrique dans son ensemble, nous avons suggéré une évaluation approfondie car les professionnels ont davantage tendance à faire le choix des partenaires en fonction d'informations dites objectives (presse spécialisée, études publiées, etc.) et en fonction des accords commerciaux qu'ils obtiennent avec les banques. Néanmoins vérifier si les cyber-consommateurs apprécient ces choix peut leur être utile.

3.2.11. Catégorie « Facilité à trouver un produit »

3.2.11.1. Trouver un produit

Tableau 88 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Trouver un produit »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Auprès d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Trouver un produit	Les internautes ont-ils facilement, rapidement trouvé le produit qu'ils recherchent ?	<p>Appréciation générale (plutôt par Q) de la facilité et/ou rapidité à trouver les produits recherchés.</p> <p>Etude auprès d'internautes expérimentés, novices et intermédiaires. Rq : exemple de caractéristiques de niveau d'expertise des internautes :</p> <ul style="list-style-type: none"> • « plus novices » (<1 an d'utilisation d'Internet), • intermédiaires (entre 1 et 4 ans), • et « plus experts » (> 4 ans). <p>Prendre aussi en compte d'autres dimensions : nombre d'heures d'utilisation d'Internet par semaine et diversité des usages (utilisation d'une messagerie, recherche d'informations, achats sur le Web, voire création de sites).</p> <p>Cf. aussi catégories « Offre-Positionnement » (la recherche de produits est aussi liée à l'offre proposée sur le site) et « Netmarketing – Présélections » (dès les premières recherches, l'offre du site correspond-elle à ce que vous vouliez ?)</p>

Le questionnaire Noël 2001 comporte une seule question interrogeant si les internautes trouvent facilement les produits qu'ils cherchent ou non, tandis que le questionnaire Noël 2002 approfondit ce point. Plusieurs questions concernent la rapidité et l'efficacité de la recherche interne par mots-clefs. Le professionnel du marketing et le responsable du Webmastering nous expliquent que la détection d'insuffisances concernant la recherche par mots-clefs a engendré la nécessité d'interroger davantage la satisfaction des internautes quant à cette dernière. Suite à cette détection, confirmée par les résultats du questionnaire Noël 2001, une activité de modification de la programmation des requêtes dans la base de « données-produits » a été entreprise. A l'époque de Noël 2002, il faut donc vérifier si les

remaniements informatiques ont porté leur fruit et si les internautes trouvent la recherche par mots-clefs plus performante. Ce point est, en effet, important ; un moteur de recherche très puissant est recommandé pour fournir les informations et produits les plus exhaustifs possibles (Lee *et al.*, 2000).

Les deux formats d'enquête, d'abord généraliste, puis plus précise, ont également été proposés dans cette catégorie. La rubrique « Trouver le produit » cherche simplement à jauger si les internautes trouvent leurs produits suffisamment rapidement et facilement, sans même distinguer s'il s'agit d'un problème d'absence des produits dans l'offre marchande ou d'un problème au niveau de la performance de la recherche. Les deux ne peuvent d'ailleurs être différenciés par l'internaute, mais une navigation difficile, empêchant ce dernier de trouver son produit, peut l'amener à se former une représentation négative de l'offre de produits (Nielsen *et al.*, 2001).

Les rubriques suivantes de cette catégorie propose d'approfondir chacune des fonctionnalités possibles dans la recherche de produits : la recherche par mot-clefs, la recherche hypertextuelle en utilisant les liens de rubriques et de sous-rubriques de produits et enfin la recherche avancée (champs multiples d'entrée de mots-clefs).

Enfin, nous avons indiqué qu'il s'agit de distinguer trois groupes de populations d'internautes (« les plus experts », « les plus novices » et « les intermédiaires »), estimés en fonction du nombre d'années d'utilisation d'Internet et du nombre d'heures de pratique par semaine, de même que dans les études du G.V.U. (Kehoe *et al.*, 1997; Kehoe *et al.*, 1998a; Kehoe *et al.*, 1998b) et de même que le réalisent les professionnels du service d'évaluation. Le responsable de la charte graphique adhère à l'utilité de cette distinction.

3.2.11.2. Recherche par mots-clefs

Tableau 89 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Recherche par mots-clefs »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Après d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Recherche par mots-clefs	Les internautes trouvent-ils que la recherche par mots-clefs leur fournit les résultats qui correspondent à leur demande ?	<p>Appréciation générale (par Q ou par E avec utilisation du site, suivis éventuellement de FG) :</p> <ul style="list-style-type: none"> de la facilité à obtenir les produits recherchés (la recherche par mots-clefs trouve les produits qui correspondent à votre recherche, toujours, souvent, parfois, rarement ?), du nombre de réponses obtenues (trop élevé, trop peu, etc.) <p>EA par E avec utilisation comparée de plusieurs sites pour mieux comprendre les difficultés rencontrées et les attentes des internautes.</p>

Cette évaluation concernant la facilité à trouver les produits par mots-clefs et la quantité de réponses obtenues est validée par les professionnels du Webmastering et du marketing et par le questionnaire Noël 2002. Les difficultés des utilisateurs à trouver le produit et la trop grande quantité de produits dans les listes de résultats sont, en effet, des problèmes avérés (Nielsen *et al.*, 2001).

Après une évaluation générale de la satisfaction des internautes, nous avons proposé une évaluation approfondie par entretiens avec utilisation comparée de différents sites marchands et/ou par *focus groups*, surtout si de fortes insatisfactions sont révélées. Les entretiens avec utilisation comparée de sites marchands choisis permettent d'obtenir des données riches sur les questions de l'utilisabilité et les *focus groups* d'approfondir les difficultés et attentes des internautes (ex. recherche par orthographe approchée au cas où ils ont oublié l'orthographe exacte d'un nom propre, exclusions de mots-clefs non pertinents pour ne pas obtenir tous les résultats avec « c » ou « d » quand ils ont entré « CD Machinchose », etc.). Nos entretiens utilisateurs et les travaux de Nielsen *et al.* (2001) montrent, en effet, qu'il existe souvent un écart entre la logique de fonctionnement de la recherche par mots-clefs (conçue par les développeurs des requêtes) et la logique de formulation d'une recherche par les utilisateurs.

3.2.11.3. Rubriques et sous-rubriques de produits

Tableau 90 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Rubriques et sous-rubriques de produits »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Après d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Rubriques et sous-rubriques de produits	<p>Les internautes trouvent-ils le choix de rubriques et sous-rubriques pertinents ?</p> <p>Selon eux, quelles rubriques manquent ?</p>	<p>Etude par Q ou par E avec utilisation du site, suivis éventuellement de FG :</p> <ul style="list-style-type: none"> • Appréciation générale de leur pertinence (les rubriques / sous-rubriques sont-elles adaptées à ce que vous cherchez, oui tout à fait, à peu près...), • Recueil d'éléments manquants, • Eventuellement, recueil de propositions de restructuration de ces rubriques et sous-rubriques, avec leurs nouveaux intitulés (question ouverte). <p>Cela complète « Page d'accueil – Rubriques et services principaux ».</p> <p>EA par E avec utilisation comparée de plusieurs sites pour mieux comprendre le découpage et les intitulés des rubriques et sous-rubriques dont les internautes ont besoin, en lien avec leur activité de recherche par liens, éventuellement associée à la recherche par mots clefs.</p> <p>Rq : Cela peut conduire à remettre en question l'architecture du site de façon plus ou moins importante.</p>

D'après nos *focus groups* essentiellement, nous avons pu observer différents comportements de recherche d'informations à partir d'un site marchand. Les observations de Nielsen (2000) concernant l'utilisation de tout site Internet corroborent les nôtres : certains internautes préfèrent utiliser la recherche par mots-clefs, d'autres préfèrent utiliser les liens et enfin il existe ceux qui utilisent les deux indifféremment, selon ce qui leur semble le plus prometteur à un moment donné. L'évaluation de la satisfaction des internautes quant à la recherche par rubriques et sous-rubriques concerne donc préférablement les internautes des seconde et troisième catégories. Son utilité est confirmée par divers travaux rendant compte de l'importance d'un bon découpage des catégories de produits, d'une bonne organisation des liens d'accès (Liu & Arnett, 2000; Nielsen *et al.*, 2001; Tilson *et al.*, 1998) et de libeller correctement ces liens (Spool *et al.*, 1999).

Cette évaluation rejoint et complète « Rubriques et services principaux » de la catégorie « Page d'accueil », puisqu'elle interroge les internautes sur l'intérêt des rubriques et sous-rubriques et donc sur leur découpage sur la page d'accueil et sur les pages suivantes. Elle propose aussi de recueillir les éventuels éléments manquants, selon l'utilisateur ou ses souhaits en termes de nouveaux découpages et intitulés de ces rubriques et sous-rubriques.

Enfin, une évaluation qualitative par entretiens avec utilisation comparée de différents sites marchands et/ou par *focus groups* figure, de même que précédemment, au registre des évaluations approfondies. D'après le professionnel de la veille-marketing et, plus particulièrement, d'après le responsable du Webmastering, ces suggestions sont pertinentes. Ce nivellement de questions de plus en plus pointues est destiné à aider les professionnels à déterminer le degré de détail nécessaire à l'évaluation et les combinaisons possibles de techniques.

3.2.11.4. Recherche avancée et autres modes de recherche

Tableau 91 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Recherche avancée »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
auprès d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Recherche avancée	<p>Les internautes trouvent-ils facilement leurs produits à l'aide de la recherche avancée ?</p> <p>Les différents champs de recherche proposés sont-ils pertinents pour leur recherche ?</p>	<p>EA pour les sites proposant cette fonctionnalité. Appréciation (par Q ou par E avec utilisation du site, suivis éventuellement de FG) :</p> <ul style="list-style-type: none"> de la facilité à obtenir les produits recherchés (la recherche avancée trouve les produits qui correspondent à votre recherche, toujours, souvent, parfois, rarement ?), de la pertinence des différents champs de recherche (lesquels utilisez-vous le plus souvent ? lesquels sont inutiles ? lesquels, actuellement manquants, seraient utiles pour vous ?). <p>EA par E avec utilisation comparée de plusieurs sites.</p>
Autres moyens de recherche	<p>Les internautes trouvent-ils utiles les autres fonctions de recherche ?</p> <p>Les trouvent-ils faciles à utiliser ?</p>	<p>Cela concerne les fonctions de recherche par fourchette de prix, par âge, par marque, par ordre alphabétique, etc.</p> <p>EA pour les sites proposant ces moyens de recherche. Appréciation (par Q ou par E avec utilisation du site) :</p> <ul style="list-style-type: none"> de l'utilité des fonctions : les internautes trouvent-ils plus facilement leurs produits grâce à ces fonctions ? S'en servent-ils ou se servent-ils plutôt des autres moyens de recherche (par mots-clefs, rubriques et sous-rubriques) ? de la facilité à les utiliser.

Cet élément a été ajouté par rapport à l'ébauche de la boîte à outils (cf. annexes du chapitre 6). Nous avons été encouragé par les propos du responsable du Webmastering et par son souci d'amélioration de la recherche simple par mots-clefs qui peut être logiquement étendu à la recherche avancée.

Il s'agit d'interroger la satisfaction des internautes quant à la facilité à obtenir des produits avec la recherche avancée. Cela concerne les internautes les plus expérimentés comme pour les plus novices : ces derniers peuvent notamment recourir à la recherche

avancée après un premier échec avec une recherche simple. Il est aussi possible d'approfondir et de savoir quels champs de recherche sont utiles, lesquels sont manquants, etc.

Deuxièmement, la possibilité d'évaluer d'autres modes de recherche (une fonction de recherche par fourchette de prix, par âge, par marque, par ordre alphabétique, etc.) a été proposée. Ces fonctions sont, en effet, très utilisées par les internautes (ex. pour trouver des idées de cadeaux) (Nielsen *et al.*, 2001; Rehman, 2000).

Ces dernières fonctionnalités n'étant pas développées par tous les sites marchands, nous avons « rangé » cette évaluation parmi les évaluations approfondies.

3.2.11.5. Listes de résultats

Tableau 92 : Outil d'aide à la conception d'évaluations de sites Internet concernant le besoin en évaluation « Listes de résultats »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Auprès d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Listes de résultats	Les internautes repèrent-ils ou choisissent-ils facilement le produit qu'ils cherchent dans les listes de résultats ?	<p>E par Q ou par E avec utilisation comparée de plusieurs sites, suivis éventuellement de FG.</p> <p>Appréciation :</p> <ul style="list-style-type: none"> de la facilité à repérer le produit dans les listes de résultats (Vous trouvez toujours très facilement votre produit, une fois parvenu aux listes de résultats, facilement, peu... ? La présence de petites photos vous aident-elles à trouver les produits plus facilement ? d'autres éléments vous aident-ils ? lesquels ?), de l'aide à trouver des produits (les présélections dans les listes de résultats vous aident-elles à trouver et choisir des produits peu précisés au départ ? ex. un guide touristique de Paris), cf. aussi rubrique « Netmarketing – présélections dans les listes de résultats », des options de hiérarchisation des produits dans les listes de résultats (quelle option est la plus utile pour vous ? : par ordre alphabétique, du moins au plus cher, du plus au moins vendu, etc.), de l'aide pour comparer les produits (ex. la possibilité de comparaison des produits dès les listes de résultats vous est-elle utile ?).
Boutons d'action – visibilité - navigation		Cf. catégorie « Look & co »

Les différents éléments ci-dessus sont aussi particulièrement appréciés par le responsable du Webmastering dont l'activité porte sur les spécifications de conception des diverses fonctionnalités et des principes de navigation sur le site.

Premièrement, il s'agit de vérifier si les internautes repèrent facilement le produit dans les listes de produits (et, si cela pose problème, de se référer à la catégorie « Look & co » pour approfondir ce point). Par exemple, dans nos entretiens utilisateurs, plusieurs internautes utilisent les photos miniatures des albums de musique ou des livres présentes dès les listes de résultats comme repère visuel, en particulier quand il ne se souviennent plus du titre. Plus généralement, les listes améliorées de résultats (avec des photos miniatures, des boutons de navigation vers la page d'accueil, vers la recherche par mots clefs, etc.) et la lisibilité de la présentation des résultats de la recherche ont, en effet, un impact important sur les internautes (Kim & Lee, 2002; Lohse & Spiller, 1998b; Tilson *et al.*, 1998).

Deuxièmement, il est possible de poser la question suivante : les présélections, dans les listes de résultats, vous aident-elles à orienter votre recherche et à choisir des produits peu précisés au départ ? Cela fait davantage partie de la recherche de produits, tout en complétant la rubrique « Présélections dans les listes de résultats – Netmarketing » (exemple de la recherche avec « guide Paris »).

Enfin, en haut des listes de résultats, il est fréquemment offert à l'internaute de choisir s'il préfère que les produits soient ordonnés par ordre alphabétique, du moins au plus cher, du plus au moins vendu, etc. Même si les professionnels du Sitemarchand.com n'évoquent pas ce point, il peut être utile de savoir si les internautes sont satisfaits de ces options et s'ils n'en manquent pas (Nielsen *et al.*, 2001). En effet, elles jouent un rôle positif dans la perception des internautes quant à la capacité du site marchand à les aider et à leur fournir un service rapide (Kim & Lee, 2002).

Il en est de même concernant la possibilité de comparer des produits à partir des listes de résultats¹⁶² (Kim & Lee, 2002). Nous avons donc proposé d'évaluer la satisfaction des internautes concernant cette fonction d'aide pour comparer les produits.

Tous ces éléments peuvent être étudiés grâce à une évaluation généraliste par questionnaire ou grâce à une évaluation qualitative, par entretiens avec utilisation comparée de différents sites et/ou par *focus groups*.

Enfin, une rubrique « Boutons d'action – visibilité – navigation » renvoie les professionnels à la catégorie « Look & co » concernant les aspects visuels de tous les éléments de la recherche.

3.2.12. Catégorie « Facilité à commander un produit »

Excepté le responsable du Webmastering, les professionnels du Sitemarchand.com sont seulement soucieux de l'utilisabilité de l'interface à partir de la « mise en panier » d'un produit et concernant tout service associé (ex. suivi de commande). Cet aspect est, à chaque fois, évoqué spontanément avant même la question de la recherche de produits. Cette focalisation est évidemment le résultat de la proportion d'internautes quittant le panier (« *On voit que sur le panier de commande il y a 50% des gens qui ne finissent pas leurs achats* ») et d'une inquiétude (« *Oui, c'est très important, ça. Ça paraît basique mais c'est vraiment fondamental pour nous. Parce que c'est notre survie* »). En effet, 40% à 80% des transactions en ligne seraient abandonnées (Nielsen *et al.*, 2001; Rehman, 2000) en raison essentiellement des difficultés rencontrées au cours du processus de commande et aussi de la lenteur de chargement des pages.

Il faut néanmoins relativiser cela. La « phase d'information » comprenant la recherche de produits et la lecture des informations les concernant est, en effet, plus déterminante que les phases d'utilisation du panier et de finalisation et réception de la commande dans le volume de ventes et le nombre de visiteurs d'un site marchand (Kim & Lee, 2002).

Dans la boîte à outils, nous avons ainsi conservé l'ordre assez logique : « Facilité à trouver un produit » puis « Facilité à commander un produit ».

3.2.12.1. Commander un produit

Tableau 93 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Commander un produit »

Élément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Après d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Commander un produit	Les internautes ont-ils facilement réussi à commander leur produit ?	Appréciation générale (plutôt par Q) de la facilité à commander des produits.

Il n'est plus besoin de démontrer l'utilité de cette évaluation... mais il faut encore distinguer trois types d'internautes, de même que pour le processus de recherche de produits. Ce premier élément cherche à savoir si les internautes, plus expérimentés, intermédiaires et plus novices, trouvent le processus pour commander un produit plutôt facile ou plutôt difficile dans son ensemble. Toujours selon le même principe, si le résultat s'avère décevant ou mitigé, il faut poursuivre l'évaluation. C'est ce que les éléments suivants permettent.

Tous les professionnels se sont montrés si soucieux de ce processus d'achat, que nous avons détaillé encore davantage d'éléments que pour la recherche. Nous nous sommes appuyé sur les données de nos entretiens utilisateurs, sur les remarques des professionnels et sur la littérature abondante sur le sujet.

3.2.12.2. Saisie des caractéristiques des produits

Tableau 94 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Saisie de la taille, de la quantité des produits »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
auprès d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Saisie des caractéristiques des produits	<p>Les internautes sélectionnent-ils facilement les caractéristiques (couleur, taille, quantité, etc.) du produit qu'ils veulent commander ?</p> <p>Ont-ils des gênes ou des difficultés, lesquelles ?</p>	<p>Appréciation générale (par Q ou par E avec utilisation d'un ou de plusieurs sites, suivis éventuellement de FG) :</p> <ul style="list-style-type: none"> • de la facilité à sélectionner le produit désiré une fois qu'il a été identifié (avez-vous trouvé facile de saisir les caractéristiques des produits que vous vouliez acheter ? très facile, facile, peu...), • Inventaire de gênes ou difficultés habituellement rencontrées, par exemple : <ul style="list-style-type: none"> ○ l'internaute novice a-t-il compris qu'il faut d'abord sélectionner les caractéristiques de produit voulu, avant de passer au panier puis au processus d'achat ? ○ quel libellé a posé problème ? ○ quel champ (ex. menu déroulant) a été difficile à saisir ? ○ l'internaute comprend-il qu'il doit cliquer sur un bouton de validation après avoir sélectionné les couleurs, tailles, etc. (si c'est le cas) ?

Nous avons d'abord proposé une évaluation générale de la satisfaction des internautes concernant la facilité de la sélection des caractéristiques du produit à commander (couleur, taille, etc.) avant ou juste après la mise en panier. Cela peut se faire par simple questionnaire.

En revanche, pour la partie suivante de l'évaluation, même s'il est possible de poser des questions ouvertes, il est préférable d'envisager une évaluation par entretiens avec observation du comportement de l'utilisateur (données qualitatives et quantitatives).

Concernant le groupe d'internautes novices, on pourra notamment vérifier s'ils comprennent qu'il faut d'abord sélectionner toutes les caractéristiques du produit voulu et qu'ensuite ils arriveront au panier (ce qui n'est pas nécessairement clair dans leur esprit).

Pour les internautes de tout niveau, il faut observer les éventuelles difficultés de compréhension des libellés des champs de saisie, les éventuelles difficultés d'utilisation des menus déroulants, des entrées libres, des boutons de validation (ex. l'internaute, après avoir

sélectionné des items dans des menus déroulants, ne voit pas qu'il faut encore cliquer sur un bouton « ajouter au panier »), etc.

3.2.12.3. Panier

Tableau 95 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Panier »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Après d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Panier	<p>Les internautes trouvent-ils le panier simple d'utilisation ?</p> <p>Ont-ils des gênes ou des difficultés, lesquelles ?</p>	<p>Etude par Q, éventuellement en ligne après abandon ou non du processus d'achat ou par E avec utilisation d'un ou de plusieurs sites, suivis éventuellement de FG. Appréciation générale :</p> <ul style="list-style-type: none"> • de la facilité à utiliser le panier (très facilement, facilement, peu...), et notamment : <ul style="list-style-type: none"> ○ de la simplicité du principe d'accès au panier (selon qu'il apparaît automatiquement à chaque fois que l'internaute met un article dans son panier ou que l'internaute lui-même doit le chercher à la fin de ses achats - un simple petit message lui indique qu'un article a été ajouté dans son panier), ○ de la simplicité du retour à la recherche d'autres produits, ○ de la simplicité de poursuite de la commande, • Inventaire de gênes ou difficultés rencontrées, par exemple : <ul style="list-style-type: none"> ○ l'internaute a-t-il remarqué qu'il peut accéder au panier depuis toutes les pages du site ? ○ l'internaute novice sait-il qu'en cliquant sur le bouton « ajouter au panier », tout le processus d'achat sera déclenché? ○ depuis le panier (voire dès la page d'accueil), l'internaute novice sait-il qu'il pourra entrer son code de bon d'achat (réduction) dans le formulaire de commande ? ○ quel champ de saisie, libellé ou message d'erreur a posé problème ? ○ l'internaute parvient-il à supprimer ou à ajouter des produits ? ○ l'internaute n'est-il pas gêné par les offres complémentaires (ex. panier trop petit, noyé au milieu des offres qui gênent sa manipulation et visibilité) ? ○ de quoi a-t-il eu peur ? qu'est-ce qui l'a bloqué ? pourquoi a-t-il abandonné le processus d'achat (si tel est le cas) ? • Inventaire d'éléments manquants gênants pour l'internaute: <ul style="list-style-type: none"> ○ l'intitulé de chaque produit sélectionné, ○ les frais d'envoi, ○ la disponibilité du produit, ○ la somme totale (cf. aussi « Récapitulatif frais & délais... »), ○ un bouton pour recalculer la nouvelle somme d'achats, etc. <p>Corrélation avec les données sur le comportement réel des visiteurs, recueillies automatiquement sur le site.</p>

Dans un premier temps, nous avons proposé de vérifier la proportion d'internautes (novices, intermédiaires et experts) satisfaits de la facilité d'utilisation du panier dans son ensemble.

Les difficultés d'utilisation observées dans nos entretiens utilisateurs et dans les travaux de Nielsen *et al.* (2001) nous ont aussi conduit à proposer d'interroger le sentiment des internautes quant à la facilité de réalisation de tâches plus précises :

- le principe d'accès au panier qui apparaît soit systématiquement à chaque article sélectionné, soit que l'internaute doit chercher (lien ou icône de panier) à la fin de ses achats (un petit message seulement indique alors à l'internaute qu'un article a été ajouté dans son panier),
- le principe de retour à la recherche d'autres produits : le bouton d'action « commander d'autres produits » renvoie généralement à la page d'accueil, au mieux à la page d'accueil de la famille de produits, mais rarement à la liste de résultats recherchés juste avant (obligeant l'internaute à refaire tout un parcours s'il veut des produits de cette même liste),
- et enfin la poursuite de la commande : notamment vérifier si les libellés des boutons d'action sont clairs (ex. confusion possible avec le libellé « poursuivre sa commande », s'agit-il de la terminer et de poursuivre d'autres « emplettes »).

Tout cela peut aussi être interrogé par questionnaire, tandis que la suite requière plutôt une évaluation par entretiens.

De même que pour la recherche de produits, nous avons proposé une évaluation par entretiens avec utilisation comparée de différents sites et/ou par *focus groups*. L'inventaire suggéré des gênes ou difficultés habituellement rencontrées à vérifier/observer est aussi bien inspiré de nos entretiens utilisateurs, des propos des professionnels que de recommandations issues de la littérature ; elles portent sur :

- la représentation des internautes, tout à fait novices, concernant le bouton « ajouter au panier » (savent-ils que le bouton déclenche tout le processus d'achat avec la saisie des coordonnées ?) et concernant le fonctionnement d'une commande avec un code de bon de réduction (savent-ils qu'ils pourront entrer leur code une fois parvenus au formulaire de commande et non dès la page d'accueil, comme ils cherchent généralement à le faire),
- l'accès au panier,
- la simplicité des champs de saisie, des libellés (Nielsen *et al.*, 2001),
- la simplicité et efficacité des messages d'erreurs (Nielsen *et al.*, 2001; Rehman, 2000),
- la visibilité du panier dans la « page du panier » : éviter les distractions trop perturbantes sur cette page ou un panier trop petit, noyé au milieu des offres qui gênent ainsi sa manipulation et visibilité (Fouchard, 2001),

- le contrôle de l'utilisateur : il doit pouvoir modifier ses choix et recalculer le montant total de son panier facilement (Fouchard, 2001; Liu & Arnett, 2000; Nielsen *et al.*, 2001),
- enfin, la présence d'une information complète sur les frais d'envoi et délais de disponibilité des produits achetés, le plus tôt possible dans le processus ; avant le passage sur le panier étant le mieux (Fouchard, 2001; Nielsen *et al.*, 2001).

Enfin, la question « de quoi l'internaute a-t-il eu peur ? », « pourquoi a-t-il abandonné le processus d'achat ? » émane des remarques des professionnels du Webmastering et du marketing qui procèdent à une analyse très fine de l'interface pour savoir pourquoi les internautes abandonnent le processus d'achat et si quelque chose les a inquiété. Pour cela, il est possible de programmer un questionnaire en ligne, aussi bien après un achat effectif qu'à la suite de tout abandon dès lors du passage par le panier.

La précédente énumération sert à pointer les gênes ou difficultés d'utilisation prépondérantes aux professionnels du Web marchand (et du Webmastering en particulier) qui n'ont pas les connaissances d'un ergonomiste, même s'ils ont une connaissance empirique des utilisateurs. De plus, ils ont tendance à se substituer à l'utilisateur, comme nous le signale le responsable de la charte graphique : « *J'avais participé à un focus avec des internautes et je trouvais ça très intéressant. Parce que soit ça nous redisait des choses qu'on savait, mais c'est toujours bon de les entendre par les autres. Et puis c'est toujours de se mettre à la place des internautes, on arrive peut-être parfois à faire des confusions, des erreurs. Et c'est vrai que c'est bien d'entendre les internautes parler eux-mêmes. De ne pas toujours se prendre... même si on est des internautes, hein* ». En effet, les Webdesigners, du fait même qu'ils sont aussi internautes, peuvent croire considérer correctement les besoins des utilisateurs, mais introduisent, en fait, des problèmes ergonomiques majeurs dans leurs productions (Chevalier & Ivory, 2003).

Les éléments de cette énumération peuvent aider tout autant à construire une évaluation adaptée, qu'à trouver des pistes d'amélioration de l'interface si nécessaire.

Pour finir, ces données peuvent être corrélées à des données objectives quantitatives recueillies sur le site (ex. quantité d'internautes ayant abandonné son achat à chaque page depuis l'arrivée sur le panier).

3.2.12.4. Bon de commande

Tableau 96 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Bon de commande »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Après d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Bon de commande	<p>Les internautes trouvent-ils le bon de commande simple à remplir ?</p> <p>Se sentent-ils en sécurité ?</p> <p>Ont-ils des gênes ou des difficultés, lesquelles ?</p>	<p>Etude par Q, éventuellement en ligne après abandon ou non du processus d'achat ou par E avec utilisation d'un ou de plusieurs sites, suivis éventuellement de FG.</p> <p>Appréciation générale :</p> <ul style="list-style-type: none"> • de la simplicité des bons de commande (où l'internaute entre ses coordonnées personnelles, ses coordonnées de paiement, le choix du mode d'expédition, etc.), • du sentiment de sécurité lors du processus d'achat, cf. catégorie « Sécurité-Paiement », • Inventaire de gênes ou difficultés rencontrées, par exemple : <ul style="list-style-type: none"> ○ quel libellé, message d'erreur, champ de saisie a posé problème ? ○ l'internaute peut-il retourner en arrière facilement ? ○ s'il est déjà identifié, ses informations personnelles sont-elles bien pré-renseignées (pour lui éviter une re-saisie) ? ○ l'organisation des champs de saisie est-elle logique (ex. selon la sémantique des informations à saisir : adresse, code postal, ville ; selon leur importance) ? ○ l'enchaînement des étapes est-il logique ? le principe de remplissage de l'adresse de facturation / adresse de livraison est-il facilement compris ? ○ le choix du mode d'expédition (relais colis, La Poste Colissimo, Chronopost, etc.) a-t-il posé problème ? ○ l'internaute est-il gêné du fait de devoir donner certaines informations personnelles ? lesquelles ? ○ pourquoi l'internaute a-t-il abandonné le processus d'achat (si tel est le cas) ?

Nous avons d'abord suggéré une évaluation générale de la satisfaction des internautes quant à la simplicité des bons de commande.

Nous avons aussi proposé un rappel concernant le sentiment de sécurité lors du processus d'achat qui renvoie les professionnels à la catégorie « Sécurité-Paiement ». Cela tient au fait que le responsable du Webmastering nous encourage à réitérer le « point sécurité » ici, puisque selon lui et d'autres aussi (Liu & Arnett, 2000), la notion de réassurance est surtout importante à ce moment.

Concernant les gênes ou difficultés habituellement rencontrées par les internautes à vérifier/observer à ce niveau, nous avons détaillé plusieurs éléments (en dehors de ceux similaires à la rubrique « Panier ») :

- vérifier si les informations concernant l'internaute sont pré-renseignées (Fouchard, 2001; Nielsen *et al.*, 2001) dans le cas où il est déjà enregistré sur le site ou identifié grâce à un *cookie*,
- vérifier, du point de vue de l'utilisateur, si l'organisation des champs de saisie est réalisée selon la sémantique des informations à saisir (adresse, code postal, ville, etc.) ou selon leur importance (champs obligatoires d'abord, puis champs facultatifs),
- vérifier si l'enchaînement des étapes est simple et rapide pour l'utilisateur (ex. le principe de remplissage de l'adresse de facturation, d'une part, et de l'adresse de livraison, d'autre part, dont la confusion peut engendrer un réel problème d'utilisabilité (Nielsen *et al.*, 2001),
- vérifier si les termes employés lors du choix du mode d'expédition n'occasionnent pas de difficultés de compréhension (que signifie « relais colis » ? Comment savoir que les termes « Rapide » et « Express » correspondent à un envoi par Colissimo (48h) et par Chronopost (24h) ?),
- vérifier si l'enchaînement des étapes pour le choix de relais colis est facilement compris,
- vérifier si le fait de devoir renseigner certaines informations personnelles n'engendre pas de blocage (nos entretiens utilisateurs et les travaux de Nielsen *et al.* (2001) montrent que les utilisateurs sont réticents à en divulguer certaines),
- comprendre pourquoi l'utilisateur a abandonné le site à ce niveau (si c'est le cas).

3.2.12.5. Récapitulatif frais & délais

Tableau 97 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Récapitulatif frais & délais »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Après d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Récapitulatif frais & délais	<p>Les internautes trouvent-ils le récapitulatif ou la facture simples à comprendre et complets ?</p> <p>Sont-ils satisfaits du message de confirmation ?</p> <p>Quels éléments sont gênants ou manquants ?</p>	<p>Etude par Q éventuellement en ligne (après abandon ou non du processus d'achat) ou par E avec utilisation d'un ou de plusieurs sites suivis éventuellement de FG.</p> <p>Appréciation générale :</p> <ul style="list-style-type: none"> • de la simplicité du récapitulatif à la fin du bon de commande ou sur la facture (reçue par courriel, voire par courrier), • de la complétude des informations données (cf. plus bas), • de l'efficacité et de la convivialité de la confirmation de commande (message de confirmation sur le site directement après la validation des bons de commande et confirmation reçue par courriel), • Inventaire d'éléments gênants et/ou manquants dans le récapitulatif ou le message de confirmation : <ul style="list-style-type: none"> ○ récapitulatif avec les frais d'expédition qui est annoncé trop tard, notamment après que l'internaute ait dû entrer ses données personnelles, ○ impossibilité d'imprimer correctement une facture ou un récapitulatif des produits achetés lors de la transaction, ○ les frais de livraison inclus dans le montant total à payer sans être précisés, en plus, séparément, ○ absence du montant total, toutes taxes comprises, ○ absence du libellé de chaque produit, ○ absence d'estimation du nombre de jours avant l'arrivée du produit, ○ absence de précision du mode d'expédition et du moyen de paiement choisis, ○ absence du numéro de dossier ou de client, si nécessaire, ○ éventuellement, absence de liens vers des informations sur les modalités de retour de produits, les contacts de divers services, etc. ou vers des possibilités de modification de la commande.

Concernant ce point, nous avons suggéré d'estimer la satisfaction des internautes quant à un ensemble d'éléments établis comme essentiels dans la littérature du domaine (Fouchard, 2001; Nielsen *et al.*, 2001; Tilson *et al.*, 1998) :

- la simplicité du récapitulatif à la fin du processus de commande ou sur la facture (reçue par courriel, voire par courrier),
- la complétude des informations fournies,
- l'efficacité et la convivialité du message de confirmation, sur le site directement après la validation des bons de commande ou reçu dans un courriel.

Enfin, un inventaire d'éléments manquants ou gênants a également été dressé, de même que précédemment : impossibilité d'imprimer une facture, récapitulatif avec les frais d'expédition annoncé après que l'internaute a dû entrer ses données personnelles, etc.

3.2.12.6. Création d'un identifiant & « compte perso »

Tableau 98 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Création d'un identifiant & Compte Perso »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Auprès d'internautes plus expérimentés, d'internautes plus novices et d'internautes intermédiaires.		
Création d'un identifiant & « compte perso »	<p>Les internautes comprennent-ils facilement le système de création d'identifiant ?</p> <p>Trouvent-ils le processus général facile d'utilisation ?</p>	<p>Rq préalable : la plupart des sites marchands demandent à l'internaute de s'identifier sur le site en entrant son adresse courriel et un mot de passe (engendrant la création d'un « compte perso »), une fois qu'il a confirmé son panier pour passer une commande.</p> <p>Appréciation générale (par Q ou par E avec utilisation d'un ou de plusieurs sites suivis éventuellement de FG) :</p> <ul style="list-style-type: none"> • de l'intérêt de cet enregistrement : Avez-vous rapidement compris l'utilité de la création d'un identifiant ou « compte perso » pour commander (ex. conservation des coordonnées d'une commande à une autre, personnalisation de la page d'accueil avec un message Bonjour Mme..., etc.) ?, • de la facilité générale d'utilisation des logins et mot de passe, • Inventaire de gênes ou difficultés habituellement rencontrées (par E), par exemple : <ul style="list-style-type: none"> ○ l'internaute comprend-il facilement comment se créer un compte ? (distinction claire des formulaires pour nouveau client et ancien client déjà enregistré), ○ les messages d'erreur ont-ils posé problème ? ○ la procédure pour retrouver son mot de passe en cas d'oubli a-t-elle posé problème ? ○ au cas où le site nécessite de se créer un compte dès la page d'accueil, quelle est la réaction de l'internaute ? <p>Cf. aussi rubrique « Compréhension de la page "compte perso" - Netmarketing».</p>
Boutons d'action – visibilité - navigation		Cf. catégorie « Look & co »

Le responsable du Webmastering nous décrit la re-conception de la précédente version du Sitemarchand.com, dont la plus importante transformation a consisté à ajouter un système de création de compte personnel (page « compte perso »). Un ensemble de recommandations a été dégagé et débattu en comité de direction, l'enjeu étant si conséquent et l'inquiétude d'une chute du taux de transformation si pesante. Cette discussion nous a conduit à étoffer encore la catégorie « Facilité à commander un produit » et a y ajouté cette rubrique « Création d'un identifiant ou "compte perso" » axée sur l'utilisabilité de cette fonctionnalité.

Nous avons proposé, une évaluation générale de la satisfaction des internautes quant à la facilité à comprendre l'utilité du procédé et quant à la facilité générale d'utilisation des logins (adresse courriel) et mots de passe.

Enfin, un inventaire des gênes ou difficultés habituellement rencontrées par l'internaute a encore été spécifié : la confusion entre formulaires d'identification pour nouveau ou ancien client, la difficulté à retrouver son mot de passe, les messages d'erreur à ce stade qui ne permettent pas de résoudre un problème et enfin la réticence des internautes face à l'éventuelle obligation de se créer un compte dès la page d'accueil. Ces éléments sont issus de divers travaux (Nielsen *et al.*, 2001; Rehman, 2000; Tilson *et al.*, 1998).

Enfin, une rubrique « Boutons d'action – visibilité – navigation » renvoie les professionnels à la catégorie « Look & co » concernant les aspects visuels de tous les éléments du processus de commande (et services associés).

3.2.13. Catégorie « Informations associées au service de vente »

3.2.13.1. Compréhension et complétude des informations associées au service de vente

Tableau 99 : Outil d'aide à la conception d'évaluations de sites Internet marchands concernant le besoin en évaluation « Compréhension et complétude des informations associées au service de vente »

Elément à évaluer	Objectif de l'évaluation	Types et détails de l'évaluation
Liste de ces informations : sur la sécurité & confidentialité, les conditions de vente, livraison & S.A.V., les contacts des services, les possibilités de personnalisation, les avantages clients, les aides d'utilisation du site.		
Pertinence & Compréhension de ces informations	<p>Les internautes trouvent-ils les informations pertinentes ?</p> <p>Comprennent-ils facilement les informations données ?</p> <p>Quelles autres formulations leur conviendraient mieux ?</p>	<p>Etude plutôt par Q ou par E. Appréciation de :</p> <ul style="list-style-type: none"> la pertinence globale de ces informations, la facilité de compréhension des informations (éventuellement de la compréhension de chaque type d'informations indépendamment sur la sécurité, les conditions de vente, la livraison...), <p>EA par E avec utilisation comparée de 2 ou 3 sites marchands (recueil de jugements comparatifs après lecture des différentes informations) pour mieux comprendre :</p> <ul style="list-style-type: none"> leurs difficultés de compréhension, leurs suggestions concernant d'autres formulations plus appropriées. <p>Rq : le contenu des informations varient selon les fonctionnalités du site marchand et les caractéristiques commerciales.</p>
Complétude de ces informations	<p>Les internautes trouvent-ils les informations assez complètes ?</p> <p>Selon eux, manque-t-il des informations importantes ?</p>	<p>Complément à ce qui précède. Etude par Q :</p> <ul style="list-style-type: none"> Appréciation de la complétude des informations (très détaillées, détaillées, pas assez, etc.). Recueil d'éventuelles informations manquantes (question ouverte). <p>EA par E avec utilisation comparée de 2 ou 3 sites marchands.</p>
Visibilité de toutes ces informations		Cf catégorie « Look & co »

Les deux professionnels du Webmastering considèrent que les pages d'explications sur les conditions de vente, la sécurité et confidentialité, les services après vente (etc.) ont besoin d'être retravaillées au niveau du contenu et du look. Vraisemblablement, ces pages ne sont pas une priorité, même si les nouvelles informations y sont intégrées « *par ci, par là* » quand nécessaire. C'est pourquoi nous avons placé cette catégorie tout à la fin dans la boîte à outils (ce qui n'était pas le cas dans l'ébauche de la boîte à outils, cf. annexes chapitre 6).

En effet, peu de sites marchands affichent des informations sur les conditions de vente, l'histoire et les références de l'entreprise (Spiller & Lohse, 1998). Pourtant, ces informations contribuent à rassurer les internautes quant à la sécurité de la transaction bancaire, à la qualité des produits et enfin quant aux garanties de livraison et de remboursement en cas de défaut (Teo, 2002).

De même que Nielsen *et al.* (2001) le préconisent, nous avons donc proposé une évaluation de la satisfaction des internautes concernant la pertinence, la facilité de compréhension et la complétude de ces informations.

Nous avons aussi indiqué une évaluation par entretiens avec utilisation comparée de différents sites et/ou par *focus groups* pour approfondir les difficultés de compréhension des internautes et recueillir leurs éventuelles suggestions de modification.

Enfin, une rubrique « Boutons d'action – Navigation - Visibilité de ces informations » renvoie les professionnels à la catégorie « Look & co » étant donné que ces aspects peuvent intervenir dans la perception et l'utilisation de ces pages.

4. CONCLUSION

Ainsi, à partir des données des entretiens de validation avec les professionnels de la veille-marketing, du marketing et du Webmastering du Sitemarchand.com, nous avons modifié l'ébauche de l'outil d'aide à la conception d'évaluations. Certains éléments des propositions d'évaluation, voire certains besoins en évaluation étaient utopistes ou inadéquats et nous les avons supprimés ou reformulés. D'autres manquaient et ont été ajoutés. Nous avons aussi légèrement modifié l'ordre global des différentes catégories pour mieux approcher les représentations pour l'action des professionnels et l'importance relative qu'ils accordent à chaque catégorie.

Cette boîte à outils s'apparente à un projet de questionnaire (Javeau, 1982). Cependant elle est plus qu'un simple projet, elle vise à aider la formulation de différents projets de questionnaire, notamment grâce aux propositions de questions déjà formulées.

Par ailleurs, la boîte à outils vise à soutenir une activité d'élaboration d'évaluations par *focus groups* ou par entretiens, avec ou sans utilisation de produits/services.

De plus, de nombreuses caractéristiques des commerces électroniques existants (fonctionnalités « online », capacités logistiques, capacités des services « offline », voire différents secteurs de vente) ont été pensées dans cette boîte à outils. Cette dernière vise, en effet, à être un outil d'aide à l'élaboration d'évaluations pour tout type de commerce électronique.

A notre sens, cette boîte à outils doit pouvoir être instrumentalisée par des professionnels de différents domaines d'activité et de différentes entreprises marchandes, selon leurs objectifs et exigences spécifiques. Nous esquissons l'activité présumée de développement d'instruments dans le chapitre de conclusion.

CHAPITRE 7

LES DEUX VERSANTS DE L'ÉVALUATION ERGONOMIQUE : SYNTHÈSE ET PERSPECTIVES

1. SYNTHÈSE ET DISCUSSIONS

Les discussions et synthèses des résultats de cette thèse sont articulées premièrement autour des éléments de connaissances sur les techniques d'élicitation d'idées par des utilisateurs de produits/systèmes (sujets de type 3), apportés à tout professionnel susceptible d'utiliser les techniques (sujet de type 1) pour produire un questionnaire et des résultats d'évaluation.

D'autre part, une discussion est menée sur la prise en compte des concepteurs et utilisateurs des résultats ergonomiques (sujets de type 2). Les particularités de l'outil d'aide à l'élaboration d'évaluations développé dans cette thèse sont ensuite explicitées au regard de recommandations concernant les méthodes d'évaluation, issues de la littérature du domaine. La prise en compte de l'activité des sujets de type 3 et l'adéquation de la boîte à outils à leurs futurs utilisateurs potentiels sont notamment mises en évidence. Les utilisateurs potentiels sont, en premier lieu, des professionnels producteurs de résultats, ayant aussi une activité de transformation des produits/services (notamment marketing ; sujets de type 1 et 2 simultanément). De plus, les professionnels de la transformation (marketing ou conception ; sujets de type 2) sont susceptibles d'utiliser la boîte à outils pour identifier leurs besoins en connaissances avant l'élaboration d'une évaluation. Ainsi, la genèse d'instruments futurs probables à partir de cet artefact (c'est-à-dire l'outil sous forme de formalisation des besoins et des possibilités d'évaluation) est imaginée pour déboucher sur l'intérêt d'une analyse plus approfondie de l'activité de sujets de type 1 et 2 avec cette boîte à outils.

Les questions du rôle de l'ergonome et du nécessaire apport réciproque des deux axes de recherche concernant l'activité des utilisateurs et celle des concepteurs sont également soulevées, ainsi que les divergences entre ergonomie et marketing.

1.1. APPORT DE CONNAISSANCES POUR LES UTILISATEURS DES TECHNIQUES D'ELICITATION

1.1.1. Concernant les dimensions utilisateurs versus experts

D'après notre étude, les données utilisateurs sont plus riches et surtout plus détaillées que les données experts. Cependant, la liste exhaustive de dimensions utilisateurs a été produite sur la base des dimensions experts, enrichies par les données utilisateurs. Nous ne pouvons

donc prétendre que les dimensions utilisateurs sont préférables aux dimensions experts pour élaborer une évaluation par questionnaire (ni l'inverse). La combinaison des deux semble souhaitable, de même que les résultats d'évaluation des sujets experts et des sujets naïfs sont généralement cumulés dans l'évaluation sensorielle des produits alimentaires.

1.1.2. Pour l'élaboration de dimensions et d'un questionnaire d'évaluation

Les résultats de l'étude comparative nous apprennent que l'entretien individuel est préférable au *focus group*, pour l'élaboration d'une liste exhaustive de dimensions concernant la perception des utilisateurs avec une grande précision, en particulier si les sujets de type 1 ont encore une connaissance pauvre du domaine (l'entretien individuel apporte des données détaillées). La meilleure productivité en termes de coûts et d'effort humain de l'entretien individuel vient aussi contredire la littérature du domaine qui, excepté certains auteurs (Morgan, 1997; Morgan & Krueger, 1993), prône le *focus group* comme la technique la plus efficiente.

Néanmoins, une étude par *focus groups* permet de dégager rapidement les principaux aspects relatifs à l'achat par Internet (les aspects commerciaux, mais aussi des aspects émotionnels et quelques aspects généraux relatifs à l'utilisabilité). Ces « méga-dimensions » correspondent à celles des questionnaires généraux les plus fréquemment utilisés dans le domaine du Web marchand (ex. SERVQUAL¹⁶³ ou les questionnaires utilisés par les professionnels du Sitemarchand.com décrits aux chapitres 5 et 6). Ainsi le *focus group* a son intérêt dans la construction d'un questionnaire interrogeant la satisfaction des utilisateurs/clients concernant plusieurs aspects à un niveau général, mais non pour la construction d'un questionnaire qui porterait sur l'utilisabilité. Dans le cas d'une étude sur l'utilisabilité des sites Internet marchands, l'entretien individuel est à nouveau préconisé puisqu'il permet d'obtenir des données précises sur ce point, en particulier si une séquence d'utilisation du produit/service est prévue.

1.1.3. Pour produire des résultats utiles à la conception du produit/service

Ce paragraphe porte exclusivement sur l'utilisation du *focus group* dans la conception. En effet, un axe de recherche à explorer davantage, de notre point de vue, est celui de la production de données à partir de *focus groups* non homogènes pour la conception. Il est généralement recommandé de mener des groupes homogènes du point de vue de ses participants (Caplan, 1990; Krueger & Casey, 2000; Maguire, 2003; Merton *et al.*, 1956; Morgan & Krueger, 1993). Or, nous suggérons de comparer des *focus groups* parfaitement homogènes à des *focus groups* présentant une certaine hétérogénéité, construite avec attention (ex. participants sensiblement de la même catégorie socioprofessionnelle mais ayant développé des schèmes d'utilisation différents). Nous pensons que ces derniers, grâce aux discussions de confrontation qu'ils génèrent (Mariampolski, 19991), contribuent particulièrement à faire émerger des idées qui ne sont pas élaborées en entretien individuel ou avec d'autres techniques d'élicitation. Ils présentent donc un intérêt jusque là peu exploré (Mariampolski confirme cela).

1.2. PRISE EN COMPTE DES PROFESSIONNELS, PRODUCTEURS ET UTILISATEURS DE RESULTATS D'EVALUATION

1.2.1. Différents mouvements en ergonomie par rapport à la prise en compte des concepteurs

Il existe un réel souci de rapprochement des communautés de l'ergonomie et de la conception (Baccino *et al.*, 2005; Jordan *et al.*, 1996b), d'une part, et de celles de l'ergonomie et du marketing, d'autre part (Dejean & Naël, 2004; Jordan *et al.*, 1996b; Leulier & Donzier, 2002). Selon nous, quatre mouvements peuvent être distingués relatifs à ces préoccupations de rapprochement et, par extension, de sensibilisation des professionnels de la transformation aux résultats d'évaluation :

- Communiquer les résultats ergonomiques,
- Faire participer les concepteurs à l'évaluation,
- Donner aux concepteurs leurs propres moyens d'évaluation,
- Prendre en compte leur activité.

1.2.1.1. Communiquer les résultats ergonomiques

Un ensemble de recommandations a, par exemple, été établi pour aider les ergonomes à décrire les problèmes d'utilisabilité afin que les concepteurs les comprennent mieux (Jeffries, 1994). Notamment, il est conseillé d'argumenter en détail les recommandations de façon à ce que les concepteurs prennent conscience des réelles incidences des problèmes sur l'efficacité dans l'utilisation du produit/système (Baccino *et al.*, 2005; Jeffries, 1994). En effet, leur compréhension n'est pas la même que celle des ergonomes et « *formuler des recommandations de façon aussi précise et ouverte que possible est délicat* » (Naël, 1989, cité par Roussel, 1996). Concernant ce point, les rapports trop longs sont aussi dénoncés. Des synthèses, schémas et illustrations sont préférables pour un meilleur travail coopératif (Baccino *et al.*, 2005; Dejean & Naël, 2004).

Néanmoins, cela constitue, sans doute, le degré le plus léger de prise en compte des concepteurs par rapport aux résultats d'évaluation, c'est-à-dire se poser la question de comment communiquer les résultats une fois qu'ils sont produits.

1.2.1.2. Faire participer les concepteurs à l'évaluation

Un degré déjà plus avancé de prise en compte des concepteurs est de les faire participer à l'évaluation, de façon à ce qu'ils soient convaincus par eux-mêmes des résultats et se les approprient directement. En effet, l'un des facteurs clefs de l'acceptation des résultats tient au fait que les concepteurs aient pu observer des sessions d'évaluation car « *seeing is believing* » (Karat, 1994, p. 216).

Une session de débriefing, impliquant évaluateurs et concepteurs, peut aussi leur permettre de débattre des recommandations envisageables après une évaluation heuristique (Nielsen, 1994a) ou après des séances de *walkthrough* (Karat, 1994). Le débriefing permet d'approfondir les causes des problèmes identifiés par les évaluateurs et de réguler les questions de langage entre évaluateurs et professionnels recevant les livrables (Karat, 1994).

Le *focus group* est, de même, une méthode favorisant la collaboration entre professionnels de différents domaines, soutenant ainsi la prise de décisions quant aux transformations à réaliser (Bruseberg & McDonagh-Philip, 2001; Langford & McDonagh, 2003a). En effet, cette technique offre l'avantage de pouvoir faire observer les *focus groups* par le client (ex. en salle de régie) ou de pouvoir faire visionner des séquences vidéo des passages les plus pertinents aux professionnels de la conception, du marketing et aux

managers. Les recommandations peuvent ainsi être acceptées plus facilement. Il est même possible de réaliser des *focus groups* avec les concepteurs, permettant ainsi qu'ils comprennent et acceptent mieux les résultats (de *focus groups* réalisés) avec des utilisateurs (Caplan, 1990).

Par ailleurs, les concepteurs sont plus enclins à utiliser le *focus group* (par exemple, en animant un eux-mêmes ; Bruseberg & McDonagh-Philp, 2001) qu'une autre méthode d'évaluation plus formelle (Bruseberg & McDonagh, 2003; Bruseberg & McDonagh-Philp, 2002), étant donné le caractère très flexible de cette méthode, permettant d'inclure divers exercices créatifs.

Dans le même ordre d'idées, une étude utilisant l'approche systémique appliquée à la qualité perçue (Nosulenko & Samoylenko, 1998, 1999b), et faisant participer concepteurs et utilisateurs, a permis de mettre au jour un écart entre les modes opératoires prévus par ces concepteurs et les opérations effectivement mises en œuvre (par eux-mêmes ou par les utilisateurs qui connaissent cependant un nombre d'échecs moins important que les concepteurs !). Cela peut aussi contribuer à faciliter la compréhension des concepteurs quant aux difficultés des utilisateurs.

1.2.1.3. Donner aux concepteurs leurs propres moyens d'évaluation

Une autre façon d'aborder la question de la prise en compte des concepteurs est de leur fournir leurs propres moyens d'évaluation. Cela n'est cependant pas sans poser des difficultés, étant donné que les concepteurs ne développent pas les mêmes compétences que les ergonomes en termes d'évaluation (Chevalier, 2005; Desurvire, 1994). Néanmoins, les méthodes d'inspection utilisant les heuristiques ou les critères ergonomiques ont été développées pour les concepteurs. L'objectif, même s'il est imparfaitement atteint, était de donner aux concepteurs la possibilité de réaliser leurs propres évaluations rapides au cours de leur activité de développement de l'interface.

Dans la même perspective, des ressources ergonomiques (check-lists, heuristiques, recommandations *et* solutions de conception associées) peuvent être rendues directement accessibles aux concepteurs (Jordan *et al.*, 1996b; Mack & Nielsen, 1993). Un ensemble de recommandations apprécié des concepteurs est le GUI¹⁶⁴ qui expose, à la fois, des principes d'utilisabilité et une boîte à outils avec des procédures de programmation déjà écrites (Baccino *et al.*, 2005). Cependant, les *guidelines* sont généralement peu appropriés aux problèmes spécifiques de la conception.

1.2.1.4. Prendre en compte l'activité des concepteurs

Le niveau de prise en compte des concepteurs le plus abouti est celui qui consiste à chercher à connaître de près leur activité. Il s'agit de la prendre en compte en amont, de façon à produire des résultats ergonomiques pertinents, afin que des transformations appropriées des produits/systèmes puissent effectivement avoir lieu. Se concentrer sur ce qui a besoin d'être produit plutôt que sur les méthodes d'évaluation, permet de débloquent des situations quand, par exemple, les professionnels sont devenus réticents à un modèle de « conception centrée utilisateurs » (Jokela, 2002). Il ne s'agit pas de modifier ni la qualité, ni le contenu de ces résultats ergonomiques mais de se centrer sur ce qu'il est utile de produire, en tenant compte des représentations des concepteurs et des exigences de leur activité. Il faut que ces résultats constituent des artefacts que ces derniers puissent instrumentaliser. Pour ce faire il est nécessaire d'étudier l'activité des utilisateurs des produits/systèmes, tout autant que celle des utilisateurs des résultats d'évaluation (Bannon & Bodker, 1991; Béguin, 1994; Folcher, 2003; Rabardel, 1995; Roussel, 1996).

1.2.1.5. La position adoptée dans cette étude

Les travaux de la seconde partie de la thèse correspondent aux deux derniers mouvements énoncés ci-dessus. Nous concevons un outil d'aide à l'évaluation pour les professionnels de la transformation et nous étudions aussi leur activité pour proposer non pas directement des résultats, mais un outil d'aide à leur production qui soit adapté à leurs besoins. Concernant le débat sur le manque de compétences des concepteurs pour évaluer des produits/systèmes, dans notre cas, l'outil s'adresse, non pas seulement à des concepteurs, mais à des professionnels, et du marketing, et de la conception, dont certains ont développé des compétences en matière d'évaluation.

Après avoir positionné notre travail par rapport aux quatre mouvements de prise en compte des destinataires des résultats, nous nous proposons de confronter l'outil conçu pour eux à ce que la littérature du domaine préconise concernant les méthodes d'évaluation.

1.2.2. Les particularités de la boîte à outils (en prenant en compte l'activité des sujets de type 1, 2 et 3)

Nous établissons d'abord une série de recommandations ou de critères concernant les méthodes d'évaluation (Bastien *et al.*, 1998; Ravden & Johnson, 1989), voire les méta-méthodes d'aide à l'élaboration d'évaluations (Hulzebosch & Jameson, 1993, 1996), sur la

base de l'examen de la littérature. Nous décrivons ensuite notre boîte à outils en rapport avec ces critères qui sont les suivants :

- Valide : les dimensions ergonomiques composant la méthode doivent se baser sur les données empiriques et les travaux théoriques qui couvrent l'ensemble du champ,
- Générale : la méthode doit être capable de s'appliquer à des contextes et des types d'interfaces variés,
- Sensible : elle doit permettre la détection d'erreurs subtiles et des comparaisons fines d'interfaces,
- Fidèle ou Systématique : la méthode doit être normalisée dans son application (la fidélité dépend de la précision des consignes adressées aux sujets, de la clarté des définitions des dimensions composant la méthode et de la formation des évaluateurs),
- Simple à utiliser (par des non-experts) : la méthode peut être utilisée par des utilisateurs des produits/systèmes (sujets de type 3), des managers (sujets de type 2), etc. ¹⁶⁵,
- Efficace : la (méta) méthode doit fournir les informations nécessaires sur la (ou les) différente(s) méthode(s)/technique(s) d'évaluation et doit faire en sorte que le recueil d'informations pour l'évaluation ne soit pas trop important,
- Participative : la méthode doit pouvoir être comprise et utilisée par un grand nombre d'« experts » ; les utilisateurs pris en tant qu'experts de leur propre utilisation (sujets type 3), les concepteurs (sujets de type 2), les ergonomes (sujets de type 1).
- Perspectives multiples : une approche modulaire dans la (méta) méthode doit permettre d'élucider les jugements d'utilisateurs tout autant que d'évaluateurs,
- Exhaustive : elle doit permettre la détection du plus grand nombre possible d'erreurs de conception (ou au minimum des problèmes les plus critiques),
- Prend en compte un usage réel du produit/système,
- « Sur mesure » (*Tailoring*) : les objectifs d'évaluation doivent être discutés avec les différents protagonistes de la situation et les résultats répondre aux intérêts de ceux qui vont s'en servir (managers, utilisateurs, etc.). De plus, la méta-méthode doit faciliter la création de listes de dimensions personnalisées pour les évaluateurs,
- Intégrée/Itérative : elle doit s'insérer au processus itératif de conception,

Notre boîte à outils répond de la façon suivante à ces critères :

- Valide :

L'étendue du champ sur lequel repose la construction de la boîte à outils est large, étant donné qu'elle se base sur les besoins en connaissances de divers professionnels (marketing,

conception, service logistique, etc.), sur la littérature du domaine et des entretiens et observations empiriques sur l'activité d'utilisation/perception (qualité perçue ergonomique, hédonique, de service, etc.).

- Générale :

La boîte à outils prend en compte une grande diversité de caractéristiques (graphiques, fonctionnelles, logistiques, commerciales, financières, etc.) de différents sites Internet marchands. En ce sens, la boîte à outils doit aider l'évaluation de tout site marchand quel que soit son format technico-commercial et quel que soit le stade d'évolution de l'entreprise dès lors qu'un commerce électronique existe (stade de fidélisation de clients existants, phase de réorientation du commerce, d'expansion, etc.). Elle permet aussi l'élaboration d'évaluations de tout site (non marchand), mais n'est pas destinée à cela au départ.

- Sensible :

Du fait de la prise en compte de la grande diversité de caractéristiques explicitées précédemment et de l'adaptabilité de la boîte à outils (voir « Sur mesure » et § 1.2.3.1.), elle doit permettre de différencier un grand nombre de sites marchands (dimensions technico-commerciales différentes) et donc de comparer avec précision leur qualité perçue.

- Fidèle – Systématique :

Au contraire de systématique, la boîte à outils est un ensemble de modules parmi lesquels, il s'agit de sélectionner ceux qui sont pertinents selon les spécificités d'une entreprise donnée, selon l'évaluation donnée à un temps « t » de la vie du site marchand. Par ailleurs, plusieurs techniques d'évaluation sont généralement proposées dans les différents volets et offrent une certaine souplesse, même si la boîte à outils est orientée vers la réalisation d'une évaluation par questionnaire.

- Simple à utiliser :

La boîte à outils est destinée à tout professionnel, un minimum sensibilisé aux techniques d'évaluation du domaine (sujets de type 1). En effet, elle est basée sur les techniques les plus utilisées dans le domaine du Web marchand, comprenant les questionnaires, les études qualitatives par entretiens ou *focus groups*, les mesures objectives d'indicateurs médiatico-commerciaux. Nous avons aussi ajouté des entretiens individuels avec mise en situation d'utilisation¹⁶⁶, même si la plupart du temps, les entreprises marchandes ont recours à des ressources extérieures si une étude ergonomique spécifique est nécessaire (ex. il n'y a pas

d'ergonome dans l'entreprise marchande de la filiale de France Télécom). Cet ajout de techniques d'évaluation ergonomiques est cependant destiné à les sensibiliser à l'intérêt de résultats ergonomiques plus approfondis (cf. § 1.2.3.2. et § 1.2.3.4.).

- Efficace :

Des explications sont fournies sur les différentes techniques évoquées ci-dessus, en préambule de la boîte à outils, qui doivent aider, si nécessaire, les professionnels. De plus, le découpage des 13 catégories et leurs intitulés doivent permettre aux sujets de type 1 (et aussi aux sujets de type 2) de comprendre facilement le contenu de chaque catégorie et de les orienter dans leur consultation de la boîte à outils. La seule entrave à son efficacité est sa longueur. La boîte à outils nécessite d'être réduite par chaque professionnel selon sa spécificité et ses besoins.

- Participative :

La boîte à outils doit pouvoir faire participer différents professionnels d'une entreprise marchande par Internet (sujets de type 1 et 2, mais pas les sujets de type 3). Elle a été conçue de façon à être simple à comprendre par divers professionnels. Une terminologie simple, assez neutre quand nécessaire (ex. « trouver le produit » plutôt que « fonctions de recherche » trop technique pour les professionnels du marketing) est employée, de même, à l'inverse, que des termes techniques du Web marchand pour qu'ils retrouvent les concepts auxquels ils sont habitués (courriels « en push », « opt-in », etc.). La boîte à outils s'adresse également aux professionnels de la logistique, du service clients et du service produits car elle contient des volets qui concernent leur activité et ses exigences.

- Perspectives multiples :

Si la boîte à outils est prévue pour faire participer différents professionnels du Web marchand (sujets de type 1 et 2), son objectif est de les aider dans la conception d'une évaluation qui concerne exclusivement le recueil de jugements des sujets de type 3. Autrement dit, les professionnels ne sont pas sollicités pour éliciter leurs propres jugements mais identifient les besoins en évaluation qui doivent être renseignés par le recueil des jugements des sujets de type 3.

Par ailleurs, la création des dimensions d'évaluation de la boîte à outils est le fruit d'un travail intégrant, la plupart du temps, des perspectives multiples. Elle repose, en effet, sur l'identification des besoins en évaluation des sujets de type 2 (conception, marketing, logistique, etc.) et souvent simultanément sur une connaissance des usages, préférences, frustrations ou difficultés d'utilisation des sujets de type 3 (cf. « Prend en compte un usage réel » et § 2.1.).

- Exhaustive :

Du fait de la grande diversité des caractéristiques considérées, explicitées précédemment, la boîte à outils doit permettre d'identifier un grand nombre de problèmes, insatisfactions ou avantages. Grâce aux listes détaillées des difficultés les plus fréquemment rencontrées par les utilisateurs, l'outil peut aussi aider à leur détection (en particulier si des entretiens individuels avec mise en situation d'utilisation sont mis en œuvre).

- Prend en compte un usage réel :

Même s'il ne s'agit pas de méthodes expérimentales ou d'observation en situation naturelle, la boîte à outils est basée sur des techniques qui permettent un recueil de données reflétant, en partie au moins, un usage réel. L'entretien avec utilisation du produit/système se déroule dans un laboratoire aménagé, suivant des scénarios d'utilisation les plus réalistes possibles. L'entretien individuel et plus particulièrement le *focus group* sont reconnus comme permettant d'obtenir des données sur les expériences vécues des participants. L'évaluation par questionnaire est sans doute la méthode la plus délicate à cet égard, surtout concernant les questions d'opinions (versus questions de fait ; Ghiglione & Matalon, 1998).

D'autre part, les dimensions d'évaluation de l'outil tiennent compte de l'activité réelle des sujets de type 3, quand c'est approprié. Par exemple, nos entretiens utilisateurs et Nielsen (Nielsen, 2000) ont identifié qu'il existe trois schèmes de recherche sur les sites Internet ou trois catégories d'utilisateurs : ceux qui préfèrent utiliser la recherche par mots-clefs, ceux qui

préfèrent utiliser les liens et ceux qui utilisent les deux selon ce qui leur semble le plus prometteur à un moment donné. Ainsi, ont été formées deux dimensions d'évaluation : celle relative à la « Recherche par mots-clef » et celle relative aux « Rubriques et sous-rubriques » qui, quant à cette dernière, tient compte de la possibilité de combiner les deux modes de recherche.

- « Sur mesure » :

L'objectif même de la conception de la boîte à outils était de partir des réels besoins en connaissances des professionnels du Web marchand et devrait faciliter un travail de discussion entre professionnels de différents domaines (cf. « Participative » et § 1.2.3.2.). De par son caractère modulaire, la boîte à outils devrait aussi faciliter une instrumentalisation spécifique par chaque professionnel selon ses objectifs (création de listes personnalisées).

- Intégrée/Itérative :

La boîte à outils peut s'insérer à différents niveaux du processus de conception. En phase amont, elle doit pouvoir être utilisée pour recueillir des informations à partir d'une ancienne version d'un site Internet marchand, en vue de la spécification d'une nouvelle. En cours de conception, elle doit pouvoir être appliquée à l'évaluation (comparative) de prototypes et/ou de solutions technico-commerciales qui vont être ainsi affinées. Enfin, un site marchand en fin de conception et/ou mis en ligne peut être évalué.

- Instrumentalisable :

Nous ajoutons une dimension à toutes les précédentes qui a déjà été suggérée, à savoir « Instrumentalisable ». Elle renvoie au fait que la boîte à outils est susceptible d'être associée à différentes fonctions (pragmatique, épistémique, réflexive) et peut être transformée dans ses composants pour devenir un instrument pour les professionnels du Web marchand.

Nous développons cette dernière dimension dans le paragraphe suivant.

1.2.3. Développement d'instruments futurs probables par les sujets de type 1 et 2 du Web marchand

Même si la boîte à outils présente une certaine forme d'aboutissement, comme tout artefact nouveau, elle connaîtra une poursuite de sa conception dans l'usage. Nous tentons ainsi d'envisager son usage futur probable et donc sa genèse instrumentale « potentielle ». Nous décrivons, tout d'abord, les aspects de l'usage prévu tel qu'inscrit dans la conception de

l'artefact (ex. son caractère à la fois exhaustif et modulaire qui se prête à une réduction selon les exigences de l'activité singulière), puis d'autres usages sentis comme prévisibles (fonctions épistémique et réflexive) d'après les entretiens avec les futurs utilisateurs potentiels de la boîte à outils, les professionnels du Web marchand.

1.2.3.1. Transformation matérielle de l'artefact en vue de former un instrument assurant une médiation pragmatique

Au contraire de systématique, la boîte à outils est très flexible et doit permettre de s'adapter à des situations d'évaluation différentes prenant en compte :

- les caractéristiques générales de l'entreprise marchande,
- l'activité et les compétences singulières de chaque professionnel,
- les objectifs d'évaluation à chaque moment où une évaluation est nécessaire.

Instrumentalisation par réduction du champ des possibles

En premier lieu, la boîte à outils devrait nécessiter d'opérer des sélections parmi l'ensemble des dimensions d'évaluation :

- il peut s'agir tout d'abord de sélectionner celles qui correspondent aux caractéristiques et besoins généraux de l'entreprise (ex. s'il n'y a pas de relais colis, inutile de conserver le besoin en évaluation « Possibilité de RDV / relais colis »),
- dans un second temps, chaque professionnel peut sélectionner les dimensions d'évaluation qui concernent plus spécifiquement son activité.

La première réduction produirait l'artefact générique à l'entreprise. Ensuite, plusieurs artefacts peuvent ainsi être générés à partir de cet artefact générique, c'est-à-dire autant qu'il y a d'utilisateurs potentiels de la boîte à outils. Ces diverses opérations de réduction doivent permettre de réduire le champ des possibles et d'adapter l'artefact à sa fonction d'identification des besoins en évaluation et d'élaboration d'évaluations, dans le contexte de l'entreprise.

Cette réduction de l'artefact correspond à une transformation matérielle réalisée par anticipation des besoins en évaluation. Dans le premier cas, il s'agit d'une transformation durable qui doit logiquement être remise en question, seulement si des modifications significatives sont réalisées dans l'entreprise marchande comme, par exemple, le développement d'un réseau de boutiques qui permette la livraison par relais colis et demande de reconsidérer cette dimension.

Dans le second cas, il s'agit aussi d'une transformation durable qui doit avoir son évolution spécifique, selon l'activité et les compétences de chaque professionnel. Une formation à la réalisation d'une nouvelle technique d'évaluation peut amener un professionnel à l'ajouter dans sa boîte à outils. De nouvelles activités dans l'entreprise peuvent amener un professionnel à considérer des dimensions anciennement écartées. Enfin, la réactivité personnelle de chaque professionnel peut aussi intervenir.

Instrumentalisation par ajout et modification de composants de l'artefact

Après cette première instrumentalisation par réduction, d'autres sont possibles par ajout ou modification de composants. En effet, il est possible que les professionnels du Web marchand, en particulier ceux responsables de la réalisation de l'évaluation (les sujets de type 1), constituent des listes spécifiques à l'entreprise et/ou personnalisées pour eux-mêmes, à partir de celles proposées dans le contenu de l'évaluation de chaque dimension (cf. colonne « Types et détails de l'évaluation » dans la boîte à outils, chapitres 5 et 6).

Il peut s'agir de listes :

- de techniques d'évaluation : en supprimant des techniques parmi celles déjà proposées, mais surtout en ajoutant celles qui sont mieux connues dans l'entreprise ou plus appréciées par le professionnel (ou encore en ajoutant d'autres techniques ergonomiques, si un service ergonomique existe dans l'entreprise),
- de formulations de questions qui peuvent être proposées dans un questionnaire : possibilités d'ajout, de suppression, de modification des questions déjà proposées,
- d'éléments intervenant dans l'activité d'utilisation/perception des sites marchands, à partir des « listes à titre indicatif » : possibilités d'ajout, de suppression, de modification de difficultés habituellement rencontrées par les internautes ou d'éléments du système/interface pertinents pour eux, selon les connaissances des professionnels quant à leur cible d'utilisateurs/clients.

Les deux premières listes doivent aider les sujets de type 1 à choisir la (ou les) technique(s) la (les) plus adéquate(s) et à formuler le questionnaire d'évaluation. La dernière liste peut aider les sujets de type 1 à être attentifs à certains éléments de l'interface lors de l'élaboration de l'évaluation (ex. création d'un scénario d'utilisation) ou peut participer à guider l'interprétation des résultats pour décider de nouvelles évaluations d'approfondissement. Par exemple, suite à une évaluation par questionnaire, la mauvaise satisfaction des utilisateurs quant à la procédure de commande peut amener les professionnels (sujets de type 1, voire des sujets de type 2) à utiliser les listes pour vérifier, d'abord par eux-

mêmes, si leur site est susceptible d'occasionner les difficultés habituellement rencontrées par les internautes. Cela peut les aider à décider de réaliser des entretiens avec mise en situation d'utilisation pour approfondir ce point et à guider la réalisation des scénarios de tâches d'utilisation.

Dans ce qui précède, la boîte à outils est destinée à s'insérer dans l'activité de construction d'une évaluation des sujets de type 1. L'instrument développé à partir de cet artefact assurerait alors une médiation pragmatique, ce pour quoi la boîte à outils est conçue en premier lieu. Selon nous, d'autres fonctions sont aussi susceptibles d'être attribuées à l'artefact, générant ainsi des instruments assurant d'autres médiations.

1.2.3.2. Attribution d'autres fonctions à la boîte à outils

D'après les entretiens avec les professionnels du Web marchand, nous pensons qu'une fonction réflexive, d'une part, et une fonction épistémique, d'autre part, sont susceptibles d'être attribuées à l'artefact.

Un instrument de type check-list pour les sujets de type 1 ou 2 (fonction réflexive)

Comme nous l'avons esquissé au début du paragraphe 1.2.3.1., la réduction de la boîte à outils par sélection de dimensions peut être réalisée en fonction de l'activité propre à chaque professionnel concernant les sujets de type 1, mais aussi concernant les sujets de type 2. En effet, ces derniers ne réalisent pas l'évaluation elle-même, mais participent généralement à l'identification des besoins en évaluation. Ils sont donc susceptibles de produire un artefact réduit selon leurs besoins en évaluation (génériques et habituels) et de l'utiliser comme check-list pour les aider dans l'identification de leurs besoins spécifiques, à chaque nouvelle évaluation.

Concrètement, lors d'une nouvelle évaluation, chaque responsable de service fait l'inventaire des derniers problèmes recensés à son niveau (ex. réclamations au service clients). Il pourrait alors utiliser sa propre boîte à outils pour vérifier s'il n'oublie rien. Dans un second temps, ces professionnels et le chargé des études se réunissent et réalisent un travail collectif pour mettre en commun les besoins précédemment identifiés. Ils pourraient alors utiliser l'artefact générique (relatif à l'entreprise dans son ensemble), comme check-list pour « faire le tour » et discuter de l'ensemble des besoins en évaluation.

Les différents artefacts peuvent donc servir à guider et contrôler l'activité individuelle ou collective d'identification des besoins en évaluation. Les propos du responsable du

Webmastering semble indiquer que la boîte à outils pourrait avoir une utilité en ce sens : « *il [l'outil] repose bien ...[...] point par point, les différents, parce qu'il y a peut-être des choses qu'on oublie un petit peu des fois* »¹⁶⁷.

Un instrument de développement de connaissances sur les possibilités d'évaluation pour les sujets de type 1 et 2 (fonction épistémique)

La boîte à outils peut aussi permettre aux professionnels (sujets de type 1, mais aussi certains sujets de type 2) de développer des connaissances sur les propriétés et potentiels d'application de chaque technique d'évaluation à partir des explications fournies en préambule ou dans le corps de l'outil. Par exemple, il est précisé, dans le besoin en évaluation « Frais d'envois - Somme minimum d'achat/livraison gratuite », que le *focus group* est particulièrement indiqué pour la recherche d'un consensus autour d'un « bon » prix. Dans chaque besoin en évaluation, il est ainsi précisé quelle est la technique la plus appropriée (en dehors du questionnaire).

Par ailleurs, la boîte à outils peut aussi permettre de faire connaître, voire accepter des dimensions écartées pour des raisons de manque de connaissances sur les techniques et possibilités de produire certains résultats. Par exemple, le responsable du Webmastering pense, d'une part, qu'il est nécessaire que les publicités en ligne soient « très voyantes » pour des raisons commerciales, d'autre part, qu'il n'est pas possible de mettre au jour les caractéristiques des publicités les plus mal acceptées, car les utilisateurs auront tendance à dire seulement ne pas en vouloir du tout. Après mise en évidence de techniques permettant de recueillir des données approfondies concernant des éléments de l'interface, même contestés, il reconnaît l'utilité de l'évaluation.

De même, nous pensons que d'autres professionnels peuvent ainsi découvrir des possibilités qui révèlent des besoins en évaluation moins évidents, mais néanmoins existants.

1.2.3.3. Instrumentation

Après avoir décrit les processus d'instrumentalisation tournés vers l'artefact, nous abordons brièvement les processus d'instrumentation et notamment l'association de la boîte à outils à un schème d'usage potentiel.

En effet, la boîte à outils est susceptible d'être associée à un schème de lecture « en entonnoir ». Le schème consisterait en un premier repérage des 13 catégories structurant la boîte à outils (liste initiale en préambule de la boîte à outils), puis en un repérage des besoins

en évaluation spécifiques (dans la première colonne de la boîte à outils, cf. chapitres 5 et 6), puis de leur objectif formulé sous la forme d'une question (dans la seconde colonne). Enfin, le contenu détaillé de l'évaluation (dans la troisième et principale colonne) est davantage susceptible d'être lu, une fois le repérage et la sélection de dimensions pertinentes réalisés.

Ce schème est relatif aux contraintes de structuration de l'action inscrites dans la structure de l'artefact, telle que pensée par le concepteur. Concernant les schèmes d'activité instrumentée orientés vers le pôle objet principal, à savoir l'évaluation à élaborer, il est nécessaire d'analyser l'activité réelle. En effet, ce schème est relatif à l'objet de l'activité et donc aux caractéristiques singulières d'une évaluation donnée, ainsi qu'à la représentation que le sujet se forge de son objet de l'activité.

1.2.3.4. Le rôle de l'ergonome, encore nécessaire ?

Pour conclure sur le développement potentiel d'instruments à partir de la boîte à outils, nous voulons encore pointer l'éventuelle nécessité de l'accompagnement de ce processus par un ergonome. Revenons sur l'exemple du responsable du Webmastering acceptant l'intérêt de résultats subjectifs sur les publicités en ligne ; la boîte à outils seule, serait-elle suffisante pour convaincre ce professionnel ? Nous ne pouvons le certifier, ce changement d'avis s'est produit au cours d'une discussion cherchant à pousser, petit à petit, dans leurs retranchements, les arguments avancés par le professionnel.

D'autre part, les différentes réductions de la boîte à outils pourraient être aidées par un ergonome, menant des entretiens avec chaque professionnel pour établir avec lui ses besoins en évaluation, puis, plus tard, pour faire le point sur les éléments positifs et négatifs de son expérience d'usage de sa boîte à outils et réfléchir à des améliorations.

Cette boîte à outils n'exclut donc pas l'intervention d'un ergonome, notamment pour contribuer à des modifications de l'artefact, mais aussi pour travailler à partir de leurs représentations pour en faire accepter certains points (notamment ergonomiques).

2. POUR CONCLURE, D'AUTRES RECHERCHES NECESSAIRES...

2.1. ETUDIER L'ACTIVITE DES UTILISATEURS DE PRODUITS/SYSTEMES ET CELLE DES CONCEPTEURS COMME DEUX AXES DE RECHERCHE COMPLEMENTAIRES

Comme nous l'avons vu, d'assez nombreux travaux sont réalisés, relatifs à comment communiquer les résultats ergonomiques, à comment faire participer les concepteurs à l'évaluation et à comment leur donner leurs propres moyens d'évaluation, mais relativement peu sont menés sur l'étude de l'activité des concepteurs pour la prendre en compte dans la production de résultats. Notre étude cherche essentiellement à contribuer à ce dernier mouvement, même si une analyse plus approfondie de l'activité des nombreux professionnels de l'entreprise marchande par Internet serait utile.

Le second axe de recherche complémentaire et indissociable du précédent est l'activité des utilisateurs des produits/systèmes.

Nous avons montré que la conception de la boîte à outils est basée sur une connaissance de l'activité des futurs utilisateurs de cet artefact (sujets de type 1 et 2), mais aussi sur une connaissance des utilisateurs des produits/systèmes (sujets de type 3).

En effet, nous avons déjà montré en quoi les dimensions d'évaluation « Publicité en ligne », « Recherche par mots-clef » et « Rubriques et sous-rubriques » émanent plus directement de l'activité des sujets de type 3 et sont reconnues (d'emblée ou non) comme des besoins en évaluation par les sujets de type 1 ou 2 du Web marchand. Inversement, une dimension d'évaluation, parmi d'autres, qui provient plus directement d'un besoin énoncé par les sujets de type 2, est « Reconnaissance visuelle ». Certes, les sujets de type 3 expriment le besoin de reconnaître l'enseigne commerciale notamment pour se sentir rassurés, mais au-delà de cela, le responsable de la charte graphique a besoin de savoir si l'identité visuelle véhiculée par le site Internet est rapidement reconnue et associée à la filiale de France Télécom. Cela constitue donc un résultat indispensable à produire du point de vue du sujet de type 2. Le résultat nécessaire à produire n'est pas le pourcentage d'utilisateurs rassurés par le logo, mais bien le pourcentage d'utilisateurs qui l'associent au Sitermarchand.com et au groupe de la filiale France Télécom.

La pertinence des dimensions d'évaluation, c'est-à-dire trouver ce qu'il est utile d'évaluer se situe donc « quelque part » entre les deux axes de recherche, posés comme

indissociables dès le début de cette thèse. Plus globalement, la pertinence d'une évaluation ne se trouve pas, de notre point de vue, dans des chemins unidirectionnels, privilégiant soit seulement les catégories très nombreuses et plus figées de l'utilisateur (cf. professionnels du service d'évaluation), soit seulement les catégories opérationnelles et plus circonstanciées des professionnels de la transformation. Un ajustement mutuel, un mouvement d'échanges et de recherche de compromis entre les deux groupes de professionnels (évaluation versus transformation) aux cultures différentes sont très souhaitables. Davantage de recherches explicitant les apports réciproques de ces deux « pôles », qui renvoient aussi à la notion de « monde commun » (Béguin, 2003, 2004b), sont encore nécessaires.

2.2. ERGONOMIE ET MARKETING

Par ailleurs, les professionnels du marketing n'ont pas la culture d'un concepteur et encore moins celle d'un ergonomiste. Notamment l'utilisabilité d'un site Internet ne fait pas partie, ni de leur culture professionnelle, ni de leurs objectifs (ce sont des aspects qui sont reconnus par eux, mais qui doivent être assurés par d'autres). Or, les professionnels du marketing sont en charge des évaluations. Ils centralisent les besoins identifiés par les différents responsables de service et font des arbitrages parmi ces besoins avec la direction. Ainsi, nous pensons, de même que Roussel (1996), que des besoins peuvent être insuffisamment pris en compte du fait que ceux qui commanditent ou dirigent les évaluations ne sont pas toujours ceux qui utilisent (tous) les résultats.

Les besoins en évaluation relatifs à la conception sont notamment moins explorés, même si des évaluations ergonomiques peuvent être ponctuellement demandées par le responsable du Webmastering. Ce dernier est, par exemple, le seul à manifester un réel souci concernant la recherche de produits sur le Sitemarchand.com. Tous les autres professionnels du marketing, mais aussi du Webmastering, ne se soucient que du panier, des formulaires de commande et des pages personnalisées du site.

Cela amène l'ergonomiste qui cherche à outiller ces professionnels à s'interroger. S'il est utile de produire des résultats que les professionnels, y compris ceux du marketing, puissent utiliser, il s'agit aussi de voir au sein de l'organisation de l'entreprise comment les divers aspects ergonomiques peuvent être davantage considérés.

Des travaux empiriques pour étudier comment l'ergonomie peut être davantage intégrée, en particulier dans les entreprises marchandes par Internet, semblent donc nécessaires.

2.3. ETUDIER L'ACTIVITE D'ELABORATION D'INSTRUMENTS A PARTIR DE L'ARTEFACT « BOITE A OUTILS »

Enfin, nous l'avons vu, il serait utile d'observer l'activité d'élaboration d'évaluations à l'aide de la boîte à outils. Il s'agirait de chercher à savoir si elle est insérée ou non comme instrument dans l'activité des professionnels du Web marchand, quels sont les processus développés pour se l'approprier et quelles sont les diverses fonctions, auxquelles elle est associée. De plus, il serait intéressant d'observer l'activité avant l'introduction de cet artefact et après, puisque « *l'accès à l'activité d'élaboration d'un instrument suppose [...] une analyse qui débute avant l'introduction d'un nouvel artefact et observe la genèse jusqu'à son aboutissement* » (Folcher, 1999).

BIBLIOGRAPHIE

- Agar, M. & MacDonald, J. (1995). Focus groups and ethnography. *Human Organization*, 54(1), 78-86.
- Amalberti, R. & Hoc, J.-M. (1998). Analyse des activités cognitives en situation dynamique : Pour quels buts ? Comment ? *Le Travail Humain*, 61(3), 209-234.
- Apple Computer (1996). Apple Web design guidelines (11 July 1996), sur le site Internet *Apple Computer*, disponible depuis: http://www.geo.tu-freiberg.de/docs/apple/web_design/intro.html.
- Artemyeva, E.Y. (1980). *Psikhlogiya subektivnoj semantiki*, (La psychologie de sémantique subjective). Moscou, M.G.U.
- Atyeo, M. & Green, R. (1994). User-friendly weapons for the competitive fight. *British Telecommunications Engineering*, 13 (October 1994), 201-205.
- Atyeo, M., Sidhu, C., Coyle, G. & Robinson, S. (1996). Working with marketing. In M.J. Tauber (Ed.) *Conference Companion on Human Factors in Computing Systems: Common Ground (CHI '96)*, Vancouver, Canada, 13-18 April 1996. New York, ACM Press, 313-314.
- Baber, C. (1996). Repertory grid theory and its application to product evaluation. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 157-165.
- Baber, C. & Stanton, N. (1996). Observation as a technique for usability evaluation. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 85-94.
- Baccino, T., Bellino, C. & Colombi, T. (2005). *Mesure de l'utilisabilité des interfaces*. Paris, Lavoisier: Hermes Sciences Publications.
- Balbo, S. (1994). *Evaluation ergonomique des interfaces utilisateur : un pas vers l'automatisation*. Thèse de doctorat, Université Grenoble 1 (Laboratoire de Génie Informatique - IMAG), Grenoble, France, septembre 1994.
- Bannon, L.J. & Bodker, S. (1991). Beyond the interface : encountering artifacts in use. In J.M. Carroll (Ed.), *Designing Interaction. Psychology at the Human-Computer*

- Interface*, Cambridge series on human-computer interaction. New York, Cambridge University Press, pp. 227-253.
- Barrett, J. & Kirk, S. (2000). Running focus groups with elderly and disabled elderly participants. *Applied Ergonomics*, 31(6), 621-629.
- Barthélémy, J. (1998). Evaluation d'une grandeur sensorielle complexe : description quantifiée. In SSHA (Ed.), *Evaluation sensorielle. Manuel méthodologique* (2^{ième} ed.). Paris, Lavoisier, coll. Tec et Doc, pp. 149-169.
- Bastien, C. & Scapin, D. (2004). La conception de logiciels interactifs centrée sur l'utilisateur : étapes et méthodes. In P. Falzon (Ed.), *Ergonomie*. Paris, PUF, pp. 451-462.
- Bastien, J.M.C. & Scapin, D.L. (1992). A validation of ergonomic criteria for the evaluation of human-computer interfaces. *International Journal of Human-Computer Interaction*, 4(2), 183-196.
- Bastien, J.M.C. & Scapin, D.L. (1993). *Ergonomic Criteria for the Evaluation of Human-Computer Interfaces* (Tech. Rep. No. 156). Rocquencourt, France, INRIA, juin 1993.
- Bastien, J.M.C. & Scapin, D.L. (1995). Evaluating a user interface with ergonomic criteria. *International Journal of Human-Computer Interaction*, 7(2), 105-121.
- Bastien, J.M.C., Scapin, D.L. & Leulier, C. (1998). Une comparaison des critères ergonomiques et des principes de dialogue ISO 9241-10 dans une tâche d'évaluation d'interface. *Revue d'Interaction Homme Machine*, 1(1), 33-63.
- Beagley, N.I. (1996). Field-based prototyping. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 95-103.
- Béguin, P. (1994). *Travailler avec la C.A.O. en ingénierie industrielle : de l'individuel au collectif dans les activités avec instruments*. Thèse de doctorat, CNAM, Laboratoire d'Ergonomie, Paris, Août 1994.
- Béguin, P. (2003). When designers meet each others in the design process. *Risky Work; the Ecologies of Risky Work with-in Complex Technological Systems*, Paris, 24-26 Juin. International Workshop.

- Béguin, P. (2004a). L'ergonome, acteur de la conception. In P. Falzon (Ed.), *Ergonomie*. Paris, PUF, pp. 375-390.
- Béguin, P. (2004b). Mondes, monde commun et versions des mondes. *Psychologie Française*, 57(1), 45-48.
- Bellman, S., Lohse, G.L. & Johnson, E.J. (1999). Predictors of online buying behavior. *Communications of the ACM*, 42(12), 32-38.
- Bevan, N. (2002). Cost-effective user-centered design. *Tutorial during the WCC2002 : IFIP 17th World Computer Congress*, Montreal, Canada, 25-30 August 2002. London, Serco Usability Services.
- Bias, R. (1991). Walkthroughs : efficient collaborative testing. *IEEE Software*, 8(5), 94-95.
- Bisseret, A., Sebillotte, S. & Falzon, P. (1999). *Techniques pratiques pour l'étude des activités expertes*. Toulouse, Octarès.
- Blanchet, A. & Gotman, A. (1992). *L'enquête et ses méthodes : l'entretien*. Paris, Nathan, coll. 128 Sociologie.
- Bodker, S., Gronbaek, K. & Kyng, M. (1993). Cooperative design: techniques and experiences from the scandinavian scene. In D. Schuler & A. Namioka (Eds.), *Participatory Design. Principles and Practices*. Hillsdale, NJ, Lawrence Erlbaum Associates, pp. 157-176.
- Bouch, A., Kuchinsky, A. & Bhatti, N. (2000). Quality is in the eye of the beholder : meeting users' requirements for internet quality of service. In T. Turner & G. Szwillus (Eds.), *The Future is Here. Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems*, The Hague, The Netherlands, 1-6 April 2000. New York, ACM Press, 297-304.
- Boy, G.A. (1996a). The group elicitation method : an introduction. In N. Shadbolt, K. O'Hara & G. Schreiber (Eds.), *Proceedings of EKAW'96: 9th European Knowledge Acquisition Workshop on Advances in Knowledge Acquisition*, Nottingham, UK, 14-17 May 1996. Lecture Notes in Computer Science. Berlin, Springer Verlag, 290-305.

- Boy, G.A. (1996b). The group elicitation method for participatory design and usability testing. In M.J. Tauber (Ed.) *Conference Companion on Human Factors in Computing Systems: Common Ground (CHI '96)*, Vancouver, Canada, 13-18 April 1996. New York, ACM Press, 87-88.
- Breivik, E. & Supphellen, M. (2003). Elicitation of product attributes in an evaluation context: A comparison of three elicitation techniques. *Journal of Economic Psychology*, 24(1), 77-98.
- Brooke, J. (1996). SUS: a 'quick and dirty' usability scale. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 189-194.
- Bruseberg, A. & McDonagh, D. (2003). Focus groups in new product development: designers' perspectives. In J. Langford, D. McDonagh & J.D. Langford (Eds.), *Focus groups: Supporting Effective Product Development*. London, Taylor & Francis, pp. 63-72.
- Bruseberg, A. & McDonagh-Philip, D. (2001). New product development by eliciting user experience and aspirations. *International Journal of Human-Computer Studies*, 55(4), 435-452.
- Bruseberg, A. & McDonagh-Philp, D. (2002). Focus groups to support the industrial/product designer : a review based on current literature and designer's feedback. *Applied Ergonomics*, 33(1), 27-38.
- Caplan, S. (1990). Using focus group methodology for ergonomic design. *Ergonomics*, 33(5), 527-533.
- Card, S.K., Moran, T.P. & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ, Lawrence Erlbaum Associates.
- Carroll, J.M. (1991). Introduction: The Kittle House Manifesto. In J.M. Carroll (Ed.), *Designing Interaction. Psychology at the Human-Computer Interface*, Cambridge series on human-computer interaction. New York, Cambridge University Press, pp. 1-16.
- Carroll, J.M. (2000). *Making Use. Scenario-Based Design of Human-Computer Interactions*. Cambridge, MIT Press.

- Carroll, J.M., Singley, M.K. & Rosson, M.B. (1992). Integrating theory development with design evaluation. *Behaviour & Information Technology*, 11(5), 247-255.
- Carton, L. (2002). Les 5 principales frustrations des internautes achetant en ligne (12 mars 2002 - n°2002-2), sur le site Internet *eShopability.com*, disponible depuis: <http://fr.eshopability.com/archives/2002-2/echiffresprint.htm>.
- Caverni, J.-P. (1988). La verbalisation comme source d'observables pour l'étude du fonctionnement cognitif. In J.P. Caverni, C. Bastien, P. Mendelsohn & G. Tiberghien (Eds.), *Psychologie cognitive : modèles et méthodes*. Grenoble, France, Presses universitaires de Grenoble, pp. 253-273.
- Chevalier, A. (2005). Evaluer un site web. Les concepteurs et utilisateurs parviennent-ils à identifier les problèmes d'utilisabilité? *Revue d'Intelligence Artificielle*, 19(1/2), 319-338.
- Chevalier, A. & Ivory, M.Y. (2003). Web site designs : influences of designer's expertise and design constraints. *International Journal of Human-Computer Studies*, 58(1), 57-87.
- Chin, J.P., Diehl, V.A. & Norman, K.L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In J.J. O'Hare (Ed.) *Proceedings of the CHI'88 Conference on Human Factors in Computing Systems*, Washington, DC, 15-19 May 1988. New York, ACM Press, 213-218.
- Clouet, G & Lambert, I (2002). Etude comparative de méthodes d'identification de la perception des usagers sur la qualité des sites Web. *Actes d'IHM'02. 14^{ème} Conférence Francophone sur l'Interaction Homme-Machine*, Poitiers, France, 26-29 novembre 2002. New York, ACM, 215-218.
- Cooper, L. & Baber, C. (2003). Focused groups: scenario-based discussions. In J. Langford, D. McDonagh & J.D. Langford (Eds.), *Focus groups: Supporting Effective Product Development*. London, Taylor & Francis, pp. 137-144.
- Coughlan, P. & Sklar, A. (2003). Bringing real world context into the focus group setting. In J. Langford, D. McDonagh & J.D. Langford (Eds.), *Focus Groups: Supporting Effective Product Development*. London, Taylor & Francis, pp. 129-135.

- Coutaz, J. (1994). Evaluation techniques : exploring the intersection of HCI and software engineering. In R.N. Taylor & J. Coutaz (Eds.), *ICSE'94 Workshop on Software Engineering and Human-Computer Interaction*, Sorrento, Italy, 16-17 May 1994. Lecture Notes in Computer Science. Berlin, Springer Verlag, 35-40.
- Crabtree, B.F., Yanoshik, M.K., Miller, W.L. & O'Connor, P.J. (1993). Selecting individual or group interviews. In D.L. Morgan (Ed.), *Successful Focus Groups. Advancing the State of the Art*. Newbury Park, CA, Sage Publications, pp. 137-149.
- Dacremont, C. (2001). Mécanismes cognitifs impliqués dans la réalisation d'épreuves triangulaires. In I. Urdapilleta, C. Ton Nu, C. Saint Denis & F. Huon de Kermadec (Eds.), *Traité d'évaluation sensorielle. Aspects cognitifs et métrologiques des perceptions*. Paris, Dunod, pp. 29-43.
- Daniellou, F. (2004). L'ergonomie dans la conduite de projets de conception de systèmes de travail. In P. Falzon (Ed.), *Ergonomie*. Paris, PUF, pp. 361-373.
- Darses, F. & Reuzeau, F. (2004). Participation des utilisateurs à la conception des systèmes et dispositifs de travail. In P. Falzon (Ed.), *Ergonomie*. Paris, PUF, pp. 405-420.
- Davis, G. (1996). Effective informal methods. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 115-119.
- Dayan, A. (2001). *Le marketing, Que sais-je ?* (10ième ed.). Paris, PUF.
- De Vries, G., Hartevelt, M. & Oosterholt, R. (1996). Private camera conversation: a new method for eliciting user responses. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 147-155.
- Dejean, P.-H. & Naël, M. (2004). Ergonomie du produit. In P. Falzon (Ed.), *Ergonomie*. Paris, PUF, pp. 463-477.
- Dellaert, B.G.C. & Kahn, B.E. (1999). How tolerable is delay?: consumers' evaluation of Internet web sites after waiting. *Journal of Interactive Marketing*, 13(1), 41-54.

- Desurvire, H.W. (1994). Faster, cheaper!! Are usability inspection methods as effective as empirical testing ? In J. Nielsen & R.L. Mack (Eds.), *Usability Inspection Methods*. New York, John Wiley & Sons, pp. 173-202.
- Dillon, A. & McKnight, C. (1990). Towards a classification of text types: a repertory grid approach. *International Journal of Man-Machine Studies*, 33(6), 623-636.
- Dix, A.J., Finlay, J.E., Abowd, G.D. & Beale, R. (1993). *Human-Computer Interaction*. New York, Prentice Hall International
- Dubois, D. (1997). Catégorisation et cognition : " 10 ans après ", une évaluation des concepts de Rosch. In D. Dubois (Ed.), *Sémantiques et cognition. Catégories, prototypes, typicalité*, Collection Sciences du langage. Paris, Edition du CNRS, pp. 31-54.
- Ede, M. (2005). Focus groups to study work practice (visité en Août 2005), sur le site Internet [Useit.com](http://www.useit.com), disponible depuis: http://www.useit.com/papers/meghan_focusgroup.html.
- Ehn, P. & Badham, R. (2002). Participatory design and the collective designer. In T. Binder, J. Gregory & I. Wagner (Eds.), *Proceedings of PDC2002, the 7th Biennial Participatory Design Conference. Inquiring Into the Politics, Contexts and Practices of Collaborative Design Work*, Malmö, Sweden, 23-25 June 2002. Palo Alto, CA, CPSR, 1-10.
- Ehn, P. & Kyng, M. (1991). Cardboard computers: mocking it-up or hands-on the future. In J. Greenbaum & M. Kyng (Eds.), *Design at Work: Cooperative Design of Computer Systems*. Hillsdale, NJ, Lawrence Erlbaum Associates, pp. 169-196.
- Ericsson, K.A. & Simon, H.A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215-251.
- Ericsson, K.A. & Simon, H.A. (1984, 1993). *Protocol Analysis : Verbal Reports as Data* (2nd ed.). Cambridge, MA, MIT Press.
- Farenc, C. (1997). *ERGOVAL: Une méthode de structuration des règles ergonomiques permettant l'évaluation automatique d'interfaces graphiques*. Thèse de doctorat, Université Toulouse 1 (LIHS), Toulouse, France, janvier 1997.

- Faulkner, L. (2003). Beyond the five-user assumption: benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments & Computers*, 35(3), 379-383.
- Fern, E.F. (1982). The use of focus groups for idea generation : the effects of group size, acquaintanceship, and moderator on response quantity and quality. *Journal of Marketing Research*, 19(1), 1-13.
- Folcher, V. (1999). *Des formes de l'activité aux formes des instruments: un exemple dans le champ du travail collectif assisté par ordinateur*. Thèse de doctorat, Université Paris 8 (UFR de Psychologie), Saint-Denis (Seine-Saint-Denis), France, Février 1999.
- Folcher, V. (2003). Appropriating artifacts as instruments: when design-for-use meets design-in-use. *Interacting with Computers*, 15(5), 647-663.
- Fouchard, G. (2001). *E-commerce dotcorps & dotcoms. La stratégie gagnante*. Paris, Osman Eyrolles Multimedia.
- Freud, S. (1954, 1999). *Cinq psychanalyses* (21^{ième} ed.). Paris, P.U.F.
- Friedrich, J. (1999). La rencontre Léontiev-Vygotski : quelques concepts clés. In Y. Clot (Ed.), *Avec Vygotski*. Paris, La Dispute/SNEDIT, pp. 167-181.
- Fuller, T.D., Edwards, J.N., Vorakitphokatorn, S. & Sermsri, S. (1993). Using focus groups to adapt survey instruments to new populations. In D.L. Morgan (Ed.), *Successful Focus Groups. Advancing the State of the Art*. Newbury Park, CA, Sage Publications, pp. 89-104.
- Gaillard, J.-P. (1997). *La psychologie de l'homme au travail. Les relations homme-machine*. Paris, Dunod.
- Gheraouti-Hélie, S. (2002). *Internet et Sécurité, Que sais-je ?* Paris, PUF.
- Ghiglione, R. & Matalon, B. (1998). *Les enquêtes sociologiques. Théories et pratique* (6^{ième} ed.). Paris, Armand Colin.
- Glass, B. (1997). Swept away in a sea of evolution : new challenges and opportunities for usability professionals. In R. Liskowsky, B.M. Velichkovsky & W. Wünschmann (Eds.), *Proceedings of the Software-Ergonomie'97. Usability Engineering: Integration von Mensch-Computer-Interaktion und Software-Entwicklung*, Dresden, Germany, 3-6

- March 1997 (German Chapter of the ACM, Volume 49). Stuttgart, Germany, B.G. Teubner, 17-26.
- Gordon, W. & Langmaid, R. (1988). *Qualitative Market Research. A Practitioner's and Buyer's Guide*. Aldershot, UK, Gower.
- Goto, K. & Cotler, E. (2002). *[Re]Design de sites Web. Conduite de Projet*. Paris, Eyrolles.
- Gould, J.D. & Lewis, C. (1985). Designing for usability : key principles and what designers think. *Communications of the ACM*, 28(3), 300-311.
- Greenbaum, J. & Madsen, K.H. (1993). Small changes: starting a participatory design process by giving participants a voice. In D. Schuler & A. Namioka (Eds.), *Participatory Design. Principles and Practices*. Hillsdale, NJ, Lawrence Erlbaum Associates, pp. 289-298.
- Greenbaum, T. (2003). The gold standard? Why the focus group deserves to be the most respected of all qualitative research tools. *Marketing Research Review*, 17(6), 22-27.
- Griffin, A. & Hauser, J.R. (1993). The voice of the customer. *Marketing Science*, 12(1), 1-27.
- Gronbaek, K., Grudin, J., Bodker, S. & Bannon, L. (1993). Achieving cooperative system design: shifting from a product to a process focus. In D. Schuler & A. Namioka (Eds.), *Participatory Design. Principles and Practices*. Hillsdale, NJ, Lawrence Erlbaum Associates, pp. 79-97.
- Haslam, R. (2003). Focus groups in health and safety research. In J. Langford, D. McDonagh & J.D. Langford (Eds.), *Focus Groups: Supporting Effective Product Development*. London, Taylor & Francis, pp. 97-113.
- Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, 13(4), 481-499.
- Hassenzahl, M., Burmester, M. & Sandweg, N. (2000a). Perceived novelty of functions - a source of hedonic quality. *Interfaces*, 42(11), 11-13.
- Hassenzahl, M., Platz, A., Burmester, M. & Lehner, K. (2000b). Hedonic and ergonomic quality aspects determine a software's appeal. In T. Turner & G. Szwillus (Eds.), *The*

- Future is Here. Proceedings of the CHI'2000 Conference on Human Factors in Computing Systems*, The Hague, The Netherlands, 1-6 April 2000. ACM Press, 201-208.
- Hassenzahl, M. & Trautmann, T. (2001). Analysis of web sites with the repertory grid technique. *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, Seattle, WA, 31 March-5 April 2001. New York, ACM Press, 167-168.
- Hassenzahl, M., Beu, A. & Burmester, M. (2001a). Engineering joy. *IEEE Software*, 18(1), 70-76.
- Hassenzahl, M., Wessler, R. & Hamborg, K.C. (2001b). Exploring and understanding product qualities that users desire. In J. Vanderdonckt, A. Blandford & A. Derycke (Eds.), *Interaction sans frontières. Actes de IHM-HCI 2001: Conférence conjointe AFIHM-BCS sur l'Interaction Homme-Machine*, Lille, France, 10-14 septembre 2001. Toulouse, France, Cepaduès Editions, 95-96.
- Helander, M. G. & Khalid, H. M. (2000). Modeling the customer in electronic commerce. *Applied Ergonomics*, 31(6), 609-619.
- Hoc, J.-M. & Leplat, J. (1983). Evaluation of different modalities of verbalization in a sorting task. *International Journal of Man-Machine Studies*, 18(3), 283-306.
- Hollnagel, E. (1999). Keep cool : the value of affective computer interfaces in a rational world. In H.-J. Bullinger & J. Ziegler (Eds.), *Proceedings of HCI'99, the 8th International Conference on Human-Computer Interaction*, Munich, Germany, 22-26 August 1999. vol.1. Mahwah, NJ, Lawrence Erlbaum Associates, 676-680.
- Holtzblatt, K. & Beyer, H. (1993). Making customer-centered design work for teams. *Communications of the ACM*, 36(10), 93-103.
- Holtzblatt, K. & Jones, S. (1993). Contextual inquiry: a participatory technique for system design. In D. Schuler & A. Namioka (Eds.), *Participatory Design. Principles and Practices*. Hillsdale, NJ, Lawrence Erlbaum Associates, pp. 177-210.
- Howard, S. & Murray, D.M. (1987). A taxonomy of evaluation techniques for HCI. In H.-J. Bullinger, B. Shackel & K. Kornwachs (Eds.), *Proceedings of INTERACT'87, the 2nd IFIP Conference on Human-Computer Interaction*, Stuttgart, Germany, 1-4 septembre 1987. Amsterdam, The Netherlands, Elsevier, 453-459.

- Hulzebosch, R. & Jameson, A. (1993). A rapid method for tailored, multiperspective evaluation of user interfaces. In S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel & T. White (Eds.), *INTERACT '93 and CHI '93 Conference Companion on Human Factors in Computing Systems*, Amsterdam, The Netherlands, 24-29 April 1993. ACM Press, 193-194.
- Hulzebosch, R. & Jameson, A. (1996). FACE : a rapid method for evaluation of user interfaces. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 195-204.
- IBM Corporation (1999). Web design guidelines, sur le site Internet *IBM*, disponible depuis: http://www-3.ibm.com/ibm/easy/eou_ext.nsf/publish/572.
- Igbaria, M., Schiffman, S.J. & Wieckowski, T.J. (1994). The respective roles of perceived usefulness and perceived fun in the acceptance of microcomputer technology. *Behaviour & Information Technology*, 13(6), 349-361.
- ISO 11 035 (1995). *Norme française ISO (International Standard Organization). Analyse sensorielle. Recherche et sélection de descripteurs pour l'élaboration d'un profil sensoriel par approche multidimensionnelle*. Association Française de Normalisation (AFNOR), juillet 1995.
- ISO 9241-10 (1996). *Norme française ISO (International Standard Organization). Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation. Principes de dialogues*. Association Française de Normalisation (AFNOR), septembre 1996.
- ISO 9241-11 (1998). *Norme française ISO (International Standard Organization). Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation. Lignes directrices concernant l'utilisabilité*. Association Française de Normalisation (AFNOR), Juin 1998.
- Ives, W. (2003). Focus groups in market research. In J. Langford, D. McDonagh & J.D. Langford (Eds.), *Focus Groups: Supporting Effective Product Development*. London, Taylor & Francis, pp. 51-62.
- Jardillier, P. (1992). *La psychologie du travail*, Que sais-je ? Paris, PUF.

- Jarvenpaa, S.L. & Todd, P.A. (1997). Consumer reactions to electronic shopping on the world wide web. *International Journal of Electronic Commerce*, 1(2), 59-88.
- Jarvenpaa, S.L., Tractinsky, N. & Vitale, M. (2000). Consumer trust in an Internet store. *Information Technology and Management*, 1, 45-71.
- Javeau, C. (1982). *L'enquête par questionnaire. Manuel à l'usage du praticien* (3^{ième} ed.). Bruxelles, Editions de l'Université de Bruxelles & Paris, Editions d'Organisation.
- Jeffries, R. (1994). Usability problem reports: helping evaluators communicate effectively with developers. In J. Nielsen & R. Mack (Eds.), *Usability Inspection Methods*. New York, John Wiley & Sons, pp. 273-294.
- Jokela, T. (2002). A method-independent process model of user-centred design. In J. Hammond, T. Gross & J. Wesson (Eds.), *Usability : Gaining a Competitive Edge. Proceedings of WCC2002 : IFIP 17th World Computer Congress*, Montreal, Canada, 25-30 August 2002. Boston, MA, Kluwer Academic Publishers, 23-38.
- Jordan, P.W., McClelland, I.L. & Thomas, B. (1996a). Introduction. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 1-3.
- Jordan, P.W., Thomas, B. & McClelland, I.L. (1996b). Issues for usability evaluation in industry: seminar discussions. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 237-224.
- Karat, C.-M. (1994). A comparison of user interface evaluation methods. In J. Nielsen & R.L. Mack (Eds.), *Usability Inspection Methods*. New York, John Wiley & Sons, pp. 203-233.
- Karat, J. (1997). User-centered software evaluation methodologies. In M. Helander, T.K. Landauer & P. Prabhu (Eds.), *Handbook of Human-Computer Interaction*. Amsterdam, The Netherlands, Elsevier Science, pp. 689-704.
- Katz, J. & Aspden, P. (1997). Motives, hurdles and dropouts. *Communications of the ACM*, 40(4), 97-102.

- Kehoe, C.M., Pitkow, J.E. & Morton, K. (1997). *GVU's eighth WWW user survey report*. Atlanta, GA, Graphics Visualization and Usability Center, Georgia Institute of Technology, December 1997. Disponible depuis: http://www.gvu.gatech.edu/user_surveys/survey-1997-10/.
- Kehoe, C.M., Pitkow, J.E. & Rogers, J.D. (1998a). *GVU's ninth WWW user survey report*. Atlanta, GA, Graphics Visualization and Usability Center, Georgia Institute of Technology, July 1998. Disponible depuis: http://www.gvu.gatech.edu/user_surveys/survey-1998-04/.
- Kehoe, C.M., Pitkow, J.E., Sutton, K., Aggarwal, G. & Rogers, J.D. (1998b). *GVU's tenth WWW user survey report*. Atlanta, GA, Graphics Visualization and Usability Center, Georgia Institute of Technology, December 1998. Disponible depuis: http://www.gvu.gatech.edu/user_surveys/survey-1998-10/.
- Kelly, G.A. (1955). *The Psychology of Personal Constructs*. New York, Norton.
- Kensing, F. & Madsen, K.H. (1993). Generating visions: future workshops and metaphorical design. In J. Greenbaum & M. Kyng (Eds.), *Design at Work: Cooperative Design of Computer Systems*. Hillsdale, NJ, Lawrence Erlbaum Associates, pp. 155-168.
- Kieras, D.E. & Polson, P. (1985). An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies*, 22, 365-394.
- Kim, J. & Lee, J. (2002). Critical design factors for successful e-commerce systems. *Behaviour & Information Technology*, 21(3), 185-199.
- Kim, J. & Moon, Y. (1998). Designing towards emotional usability in customer interfaces - trustworthiness of cyber-banking system interfaces. *Interacting With Computers*, 10(1), 1-29.
- Kim, J. & Yoo, B. (2000). Toward the optimal link structure of the cyber shopping mall. *International Journal of Human-Computer Studies*, 52(3), 531-551.
- Kirakowski, J. (1996). The software usability measurement inventory : background and usage. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability evaluation in industry*. London, Taylor & Francis, pp. 169-177.

- Kirakowski, J. & Claridge, N. (2001). A professional tool for evaluating Web sites. In J. Vanderdonckt, A. Blandford & A. Derycke (Eds.), *Interactions sans frontières. Actes de IHM-HCI 2001: Conférence conjointe AFIHM-BCS sur l'Interaction Homme-Machine*, Lille, France, 10-14 septembre 2001. Toulouse, France, Cépaduès Editions, 93-94.
- Kirakowski, J., Claridge, N. & Whitehand, R. (1998). Human centered measures of success in Web site design. *Virtual Proceedings of the 4th Conference on Human Factors and the Web*, 5 June 1998. Basking Ridge, NJ, AT&T Labs. Disponible depuis: <http://www.research.att.com/conf/hfweb/proceedings/kirakowski/index.html>.
- Kirakowski, J. & Corbett, M. (1988). Measuring user satisfaction. In D.M. Jones & R. Winder (Eds.), *People and Computers IV. Proceedings of the HCI'88*, Manchester, UK. BCS Conference Series. Cambridge, UK, Cambridge University Press, 329-338.
- Kitzinger, J. (1994). Focus groups : method or madness ? In M. Boulton (Ed.), *Challenge and Innovation. Methodological Advances in Social Research on HIV-AIDS*. London, Taylor & Francis, pp. 159-175.
- Konradt, U., Wandke, H., Balazs, B. & Christophersen, T. (2003). Usability in online shops: scale construction, validation and the influence on the buyers' intention and decision. *Behaviour & Information Technology*, 22(3), 165-174.
- Köster, E.P. (1998a). L'organisation des épreuves hédoniques. In SSHA (Ed.), *Evaluation sensorielle. Manuel méthodologique* (2^{ième} ed.). Paris, Lavoisier, coll. Tec et Doc, pp. 181-184.
- Köster, E.P. (1998b). Les épreuves. In SSHA (Ed.), *Evaluation sensorielle. Manuel méthodologique* (2^{ième} ed.). Paris, Lavoisier, coll. Tec et Doc, pp. 185-206.
- Krueger, R.A. & Casey, M.A. (2000). *Focus Group : A Practical Guide for Applied Research* (3rd ed.). Thousand Oaks, CA, Sage Publications.
- Kuhn, K. (2000). Problems and benefits of requirements gathering with focus groups: a case study. *International Journal of Human-Computer Interaction*, 12(3&4), 309-325.
- Kwon, O.B., Kim, C.-R. & Lee, E.J. (2002). Impact of website information design factors on consumer ratings of web-based auction sites. *Behaviour & Information Technology*, 21(6), 387-402.

- Lahlou, S., Nosulenko, V. & Samoylenko, E. (2002). Un cadre méthodologique pour le design des environnements augmentés. *Informations sur les Sciences Sociales*, 41(4), 471-530.
- Lambert, I. & Legrand, M. (2001). *Analyse des remontées clientèle. Comment les internautes perçoivent-ils le CKM (customer knowledge management) ?* (Rapport de FTRD/DIH/EQS). Rennes, FranceTélécom R&D, septembre/octobre 2001.
- Langford, J. & McDonagh, D. (2003a). Introduction. In J. Langford, D. McDonagh & J.D. Langford (Eds.), *Focus Groups: Supporting Effective Product Development*. London, Taylor & Francis, pp. 1-18.
- Langford, J. & McDonagh, D. (2003b). Focus group tools. In J. Langford, D. McDonagh & J.D. Langford (Eds.), *Focus Groups: Supporting Effective Product Development*. London, Taylor & Francis, pp. 173-224.
- Langford, J., Wilson, J.R. & Haines, H. (2003). Beyond focus groups? The use of group discussion methods in participatory design. In J. Langford, D. McDonagh & J.D. Langford (Eds.), *Focus Groups: Supporting Effective Product Development*. London, Taylor & Francis, pp. 157-170.
- Le Calonnec, M. (2001). *Méthodologies d'enquêtes* (Rapport de FTRD/DIH/EQS). Rennes, FranceTélécom R&D, 12 avril 2001.
- Lee, J., Kim, J. & Moon, Y. (2000). What makes internet users visit cyber stores again ? Key design factors for customer loyalty. In T. Turner & G. Szwillus (Eds.), *The Future is Here. Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems*, The Hague, The Netherlands, 1-6 April 2000. New York, ACM Press, 305-312.
- Leplat, J. & Hoc, J.-M. (1981). Subsequent verbalization in the study of cognitive processes. *Ergonomics*, 24(10), 743-755.
- Les_Cahiers_du_Club_CRIN (1997). *Evaluation subjective. Méthodes, applications et enjeux*. Paris, Association ECRIN.
- Leulier, C, Bastien, J.M.C. & Scapin, D.L. (1998). *Commerce & Interactions. Compilation of Ergonomic Guidelines for the Design and Evaluation of Web Sites* (Web Guidelines Rep. n° EP 22287). Rocquencourt, France, INRIA, February 1998.

- Leulier, C, Bastien, J.M.C. & Scapin, D.L. (1999). L'évaluation des sites web : apports respectifs des questionnaires de satisfaction et de l'inspection ergonomique. In J. Nanard & P. Girard (Eds.), *Interaction pour tous. Actes d'IHM'99, 11^{ème} Conférence Internationale Francophone sur l'Interaction Homme-Machine*, Montpellier, France, 22-26 novembre 1999. Toulouse, France, Cépaduès Editions, 49-52.
- Leulier, C & Donzier, P-A. (2002). Ergonomie et marketing : deux disciplines complémentaires pour des sites Web adaptés aux utilisateurs. *L'Homme et les Nouvelles Technologies de l'Information et de la Communication. Usages et Usagers. Actes de la Conférence ERGO-IA 2002*, Biarritz, 8-10 octobre 2002. Bidart, France, Estia & Estia Innovation, 159-175.
- Leventhal, L., Teasley, B., Blumenthal, B., Instone, K., Stone, D. & Donskoy, M.V. (1996). Assessing user interfaces for diverse user groups : evaluation strategies and defining characteristics. *Behaviour & Information Technology*, 15(3), 127-137.
- Levy-Leboyer, C., Gosse, M. & Naturel, V. (1985). Une nouvelle vieille méthode : la Repertory Grid de Kelly. *Revue de Psychologie Appliquée*, 35(4), 255-269.
- Lewis, J.R. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASK. *SIGCHI Bulletin*, 23(1), 78-81.
- Lewis, J.R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1), 57-78.
- Lin, H.X., Choong, Y.-Y. & Salvendy, G. (1997). A proposed index of usability : a method for comparing the relative usability of different software systems. *Behaviour & Information Technology*, 16(4/5), 267-278.
- Lincoln, Y.S. & Guba, E.G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA, Sage Publications.
- Lindgaard, G. (2002). Deconstructing silos : the business value of usability in the 21st century. In J. Hammond, T. Gross & J. Wesson (Eds.), *Usability : Gaining a Competitive Edge. Proceedings of WCC2002 : IFIP 17th World Computer Congress*, Montreal, Canada, 25-30 August 2002. Boston, MA, Kluwer Academic Publishers, 3-20.

- Lindgaard, G. & Dudek, C. (2002). User satisfaction, aesthetics and usability: beyond reductionism. In J. Hammond, T. Gross & J. Wesson (Eds.), *Usability : Gaining a Competitive Edge. Proceedings of WCC2002 : IFIP 17th World Computer Congress*, Montreal, Canada, 25-30 August 2002. Boston, MA, Kluwer Academic Publishers, 231-246.
- Liu, C. & Arnett, K.P. (2000). Exploring the factors associated with Web site success in the context of electronic commerce. *Information and Management*, 38(1), 23-33.
- Loftus, G.R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, 3(4), 525-551.
- Lohse, G.L., Bellman, S. & Johnson, E.J. (2000). Consumer buying behavior on the Internet : findings from panel data. *Journal of Interactive Marketing*, 14(1), 15-29.
- Lohse, G.L. & Spiller, P. (1998a). Electronic shopping. *Communications of the ACM*, 41(7), 81-87.
- Lohse, G.L. & Spiller, P. (1998b). Quantifying the effect of user interface design features on cyberstore traffic and sales. In C.-M. Karat, A. Lund, J. Coutaz & J. Karat (Eds.), *Proceedings of the CHI'98 Conference on Human Factors in Computing Systems*, Los Angeles, CA, 18-23 April 1998. New York, ACM Press, 211-218.
- Lomov, B.F. (1975). Sistemnyj podkhod v inzhenernoj psikhologii (L'approche systémique de la psychologie de l'ingénierie). *Studia Psychologica*, 17(2), 86-93.
- Lomov, B.F. (1977). O putyakh postroeniya teorii inzhenernoj psikhologii na osnove sistemnogo podkhoda (Les perspectives de la conception d'une théorie de la psychologie de l'ingénierie dans le cadre de l'approche systémique). In B.F. Lomov, V.F. Rubakhin & V.F. Venda (Eds.), *Inzhenernaya psikhologiya*, (La psychologie de l'ingénierie). Moscou, Nauka, pp. 31-54.
- Lomov, B.F. (1984). *Metodologitcheskiye i teoretitcheskiye problemy psikhologii*, (Les problèmes méthodologiques et théoriques en psychologie). Mosou, Nauka.
- Mac Leod, P., Sauvageot, F. & Köster, E.P. (1998). Les caractéristiques d'une réponse sensorielle. In SSHA (Ed.), *Evaluation sensorielle. Manuel méthodologique* (2^{ième} ed.). Paris, Lavoisier, coll. Tec et Doc, pp. 5-30.

- Macarez, N. & Leslé, F. (2001). *Le commerce électronique*, Que sais-je ? Paris, PUF.
- Macaulay, L.A. (1996). *Requirements Engineering*, Series on Applied Computing. London, Springer Verlag.
- Macbeth, S.A., Moroney, W.F. & Biers, D.W. (2000). Development and evaluation of symbols and icons : a comparison of the production and focus group methods. *Ergonomics for the New Millennium. Proceedings of IEA2000 / HFES2000 : 14th Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Society*, San Diego, CA, 29 July - 4 August 2000. Santa Monica, CA, Human Factors and Ergonomics Society, 327-329.
- Mack, R. & Nielsen, J. (1993). Usability inspection methods: report on a workshop held at CHI'92, Monterey, CA, 3-4 May, 1992. *SIGCHI Bulletin*, 25(1), 28-33.
- Macleod, M. (1996). Performance measurement and ecological validity. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 227-235.
- Macleod, M., Bowden, R., Bevan, N. & Curson, I. (1997). The MUSiC performance measurement method. *Behavior & Information Technology*, 16(4/5), 279-293.
- Madsen, K.H. & Aiken, P.H (1993). Experiences using cooperative interactive storyboard prototyping. *Communications of the ACM*, 36(6), 57-64.
- Maguire, M. (2003). The use of focus groups for user requirements analysis. In J. Langford, D. McDonagh & J.D. Langford (Eds.), *Focus Groups: Supporting Effective Product Development*. London, Taylor & Francis, pp. 73-96.
- Maguire, M. & Bevan, N. (2002). User requirements analysis: a review of supporting methods. In J. Hammond, T. Gross & J. Wesson (Eds.), *Usability : Gaining a Competitive Edge. Proceedings of WCC2002 : IFIP 17th World Computer Congress*, Montreal, Canada, 25-30 August 2002. Boston, MA, Kluwer Academic Publishers, 133-148.
- Manchester_Institute_for_Popular_Culture (1999). *Key issues in web-site usability* (Report Version 3.0). Manchester Institute for Popular Culture, September 1999.

- Marchadour, A-G. (2001). *Modalités de passation d'un questionnaire* (Rapport de FTRD/DIH/EQS). Rennes, FranceTélécom R&D, 20 avril 2001.
- Mariampolski, H. (1991). Beyond conventional focus groups : emerging options for qualitative research. *Marketing Research Review*, 5(10), 16-43.
- Marshall, C., Nelson, C. & Gardiner, M.M. (1987). Design guidelines. In M.M. Gardiner & B. Christie (Eds.), *Applying Cognitive Psychology to User-Interface Design*. Chichester, UK, John Wiley & Sons, pp. 221-278.
- Mayhew, D.J. (1999). *The Usability Engineering Lifecycle. A Practitioner's Handbook for User Interface Design*. San Francisco, CA, Morgan Kaufmann Publishers.
- Merton, R.K., Fiske, M. & Kendall, P.L. (1956). *The Focused Interview, A Manual of Problems and Procedures*. Glencoe, IL, Free Press.
- Merton, R.K. & Kendall, P.L. (1946). The focused interview. *American Journal of Sociology*, 51, 541-557.
- Molich, R. & Nielsen, J. (1990). Improving a human-computer dialogue. *Communications of the ACM*, 33(3), 338-348.
- Mollo, V. & Falzon, P. (2004). Auto- and allo-confrontation as tools for reflective activities. *Applied Ergonomics*, 35(6), 531-540.
- Morgan, D.L. (1993). Future directions for focus groups. In D.L. Morgan (Ed.), *Successful Focus Groups. Advancing the State of the Art*. Newbury Park, CA, Sage Publications, pp. 225-244.
- Morgan, D.L. (1997). *Focus Group as Qualitative Research*, Qualitative Research Methods Series, vol. 16 (2nd ed.). Thousand Oaks, CA, Sage Publications.
- Morgan, D.L. & Krueger, R.A. (1993). When to use focus groups and why. In D.L. Morgan (Ed.), *Successful Focus Groups. Advancing the State of the Art*. Newbury Park, CA, Sage Publications, pp. 3-19.
- Mucchielli, R. (1968). *Le questionnaire dans l'enquête psycho-sociale*, Connaissance du problème, applications pratiques. Séminaire de Roger Mucchielli. Paris, Librairies techniques, Editions sociales françaises.

- Mucchielli, R. (1996). *L'interview de groupe* (7^{ème} ed.). Paris, ESF éditeur.
- Muller, M.J. (1997). Participatory practices in the software lifecycle. In M. Helander, T.K. Landauer & P. Prabhu (Eds.), *Handbook of Human-Computer Interaction*. Amsterdam, The Netherlands, Elsevier Science, pp. 255-297.
- Muller, M.J., Tudor, L.G., Wildman, D.M., White, E.A., Root, R.W., Dayton, T., Carr, R., Diekmann, B. & Dykstra-Erickson, E.A. (1995). Bifocal tools for scenarios and representations in participatory activities with users. In J.M. Carroll (Ed.), *Scenario-Based Design. Envisioning Work and Technology in System Development*. New-York, Wiley, pp. 19-36.
- Naël, M. (1989). Rapport de synthèse sur le thème "ergonomie du produit". XXV^{ième} congrès de la SELF. *Evolutions technologiques et Ergonomie*, Lyon, France, 3-6 octobre 1989. Paris, PUF.
- Nielsen, J. (1992). Finding usability problems through heuristic evaluation. In P. Bowersfeld, J. Bennett & G. Lynch (Eds.), *Striking a Balance. Proceedings of the CHI'92 Conference on Human Factors in Computing Systems*, Monterey, CA, 3-7 May 1992. New York, ACM Press, 373-380.
- Nielsen, J. (1993). *Usability Engineering*. Boston, MA, Academic Press.
- Nielsen, J. (1994a). Heuristic Evaluations. In J. Nielsen & R.L. Mack (Eds.), *Usability Inspection Methods*. New York, John Wiley & Sons, pp. 25-61.
- Nielsen, J. (1994b). Enhancing the explanatory power of usability heuristics. In B. Adelson, S. Dumais & J. Olson (Eds.), *Celebrating Interdependence. Proceedings of the CHI'94 Conference on Human Factors in Computing Systems*, Boston, MA, 24-28 April 1994. Vol. 1. New York, ACM Press, 152-158.
- Nielsen, J. (1996). Top ten mistakes in web design, sur le site Internet *Useit.com*, disponible depuis: www.useit.com/alterbox/9605.html.
- Nielsen, J. (1997). The use and misuse of focus groups, sur le site Internet *Useit.com*, disponible depuis: <http://www.useit.com/papers/focusgroups.html>.
- Nielsen, J. (1998). Estimated cost of running a focus group, sur le site Internet *Useit.com*, disponible depuis: <http://www.useit.com/papers/focusgroupcost.html>.

- Nielsen, J. (1999). Voodoo usability. Alertbox, December 12, 1999, sur le site Internet *Useit.com*, disponible depuis: <http://www.useit.com/alertbox/991212.html>.
- Nielsen, J. (2000). *Conception de sites Web. L'art de la simplicité*. Paris, CampusPress France.
- Nielsen, J. & Landauer, T.K. (1993). A mathematical model of the finding of usability problems. In B. Arnold, G. Van der Veer & T. White (Eds.), *Proceedings of the INTERACT'93 and CHI '93 Conference on Human Factors in Computing Systems*, Amsterdam, The Netherlands, 24-29 April 1993. New York, ACM Press, 206-213.
- Nielsen, J. & Molich, R. (1990). Heuristic evaluation of user interfaces. In J.C. Chew & J. Whiteside (Eds.), *Empowering People. Proceedings of the CHI'90 Conference on Human Factors in Computing Systems*, Seattle, WA, 1-5 April 1990. New York, ACM Press, 249-256.
- Nielsen, J., Mollich, R., Snyder, C. & Farrell, S. (2001). *E-commerce User Experience*. Nielsen Norman Group.
- Nisbett, R.E. & Wilson, T.D. (1977). Telling more than we can know. Verbal reports on mental processes. *Psychological Review*, 84(3), 231-259.
- Nogier, J.-F. (2002). *De l'ergonomie du logiciel au design des sites web*. Paris, Dunod, coll. Informatiques.
- Nosulenko, V., Parizet, E. & Samoylenko, E. (1998). La méthode d'analyse des verbalisations libres : une application à la caractérisation des bruits de véhicules. *Informations sur les Sciences Sociales*, 37(4), 593-611.
- Nosulenko, V., Parizet, E. & Samoylenko, E. (2000). Différences individuelles de perception de bruits de véhicules à moteur diesel. *Revue Française du Marketing*, 4/5(179/180), 157-168.
- Nosulenko, V. & Samoylenko, E. (1997). Approche systémique de l'analyse des verbalisations dans le cadre de l'étude des processus perceptifs et cognitifs. *Informations sur les Sciences Sociales*, 36(2), 223-261.
- Nosulenko, V. & Samoylenko, E. (1998). *Analyse de la qualité perçue des produits I.S.P.* (Contrat n°981B074 entre le CNET et l'ANVIE). Paris, ANVIE (Association

Nationale pour la Valorisation Interdisciplinaire de la recherche en sciences de l'homme et de la société auprès des Entreprises).

Nosulenko, V. & Samoylenko, E. (1999a). *Synthèse des fondements de l'approche élaborée pour l'étude de la qualité perçue des produits et services* (Lot 1 du contrat n° 991B169 entre le CNET et l'ANVIE). Paris, ANVIE (Association Nationale pour la Valorisation Interdisciplinaire de la recherche en sciences de l'homme et de la société auprès des Entreprises).

Nosulenko, V. & Samoylenko, E. (1999b). Evaluation de la qualité perçue des I.S.P. par verbalisations. *Ergonomie et Télécommunications. Journée satellite du XXXIV^{ème} Congrès de la SELF*, Caen, France, 14 septembre. France Telecom, 99-105.

Nosulenko, V. & Samoylenko, E. (2000a). *Analyse de la qualité perçue des produits I.S.P. Evaluation comparative de Wanadoo, Club Internet et AOL* (Lot 2 du contrat n°991B169 entre le CNET et l'ANVIE). Paris, ANVIE (Association Nationale pour la Valorisation Interdisciplinaire de la recherche en sciences de l'homme et de la société auprès des Entreprises).

Nosulenko, V. & Samoylenko, E. (2000b). *Analyse de la qualité perçue des produits de visiophonie. Evaluation comparative de NetMeeting, CU-SeeMe et CnetConferencing* (Lot 3 du contrat n°991B169 entre le CNET et l'ANVIE). Paris, ANVIE (Association Nationale pour la Valorisation Interdisciplinaire de la recherche en sciences de l'homme et de la société auprès des Entreprises).

Nosulenko, V. & Samoylenko, E. (2001). Evaluation de la qualité perçue des produits et services : approche interdisciplinaire. *International Journal of Design and Innovation Research*, 2(2), 35-60.

Nosulenko, V., Samoylenko, E. & Parizet, E. (1996). Evaluation and verbal comparison of noises produced by car engines. *International Journal of Psychology*, 31(3-4), 51 (Abstracts of the XXVI International Congress of Psychology, Montreal, Canada, 16-21 August 1996).

O'Brien, K. (1993). Improving survey questionnaires through focus groups. In D.L. Morgan (Ed.), *Successful Focus Groups. Advancing the State of the Art*. Newbury Park, CA, Sage Publications, pp. 105-117.

- O'Donnell, P.J., Scobie, G. & Baxter, I. (1991). The use of focus groups as an evaluation technique in HCI. In D. Diaper & N. Hammond (Eds.), *People and Computers VI. Proceedings of the HCI'91*, Edinburgh, UK, 20-23 August 1991. BCS Conference Series. Cambridge, UK, Cambridge University Press, 211-224.
- Parasuraman, A., Berry, L.L. & Zeithaml, V.A. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, 49 (Fall 1985), 41-50.
- Parasuraman, A., Berry, L.L. & Zeithaml, V.A. (1991). Refinement and reassessment of the SERVQUAL scale. *Journal of Retailing*, 67(4), 420-450.
- Parasuraman, A., Berry, L.L. & Zeithaml, V.A. (1993). More on improving service quality measurement. *Journal of Retailing*, 69(1), 140-147.
- Park, S., Choi, D. & Kim, J. (2004). Critical factors for the aesthetic fidelity of web pages: empirical studies with professional web designers and users. *Interacting with Computers*, 16(2), 351-376.
- Piaget, J. (1974). *La prise de conscience*. Paris, P.U.F.
- Pican, N. (2001). *Application de Mhybic aux portails Web généralistes - Méthode hybride d'identification de critères qualité* (Rapport de FTRD/DIH/EQS). Rennes, FranceTélécom R&D, 23 juillet 2001.
- Pilverdier, S.L. & Gaillard, F. (2001). Etude des principales difficultés rencontrées par les acheteurs en ligne (Newsletter n°51, Atelier BNP-Paris Bas du 26 Juin 2001), sur le site Internet *Axance*, disponible depuis: http://www.axance.com/00download/02etudes/axance_ecommerc_juin01.ppt.
- Pitkow, J.E. & Kehoe, C.M. (1996). Emerging trends in the WWW user population. *Communications of the ACM*, 39(6), 106-108.
- Polson, P.G., Lewis, C., Rieman, J. & Wharton, C. (1992). Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies*, 36(5), 741-773.
- Rabardel, P. (1995). *Les hommes et les technologies. Approche cognitive des instruments contemporains*. Paris, Armand Colins.

- Ravden, S.J. & Johnson, G.I. (1989). *Evaluating Usability of Human-Computer Interfaces. A Practical Method*, Ellis Horwood books in information technology. Chichester, UK, Ellis Horwood.
- Rehman, A. (2000). *Holiday 2000 E-Commerce*. New York, Creative Good, October 2000. Disponible depuis: www.creativegood.com.
- Rijken, D. & Mulder, B. (1996). Information ecologies, experience and ergonomics. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 49-57.
- Roussel, B. (1996). *Ergonomie en conception de produits. Proposition d'une méthode centrée sur la formulation de principes de solutions ergonomiques dans le processus interdisciplinaire de conception de produits*. Thèse de doctorat, ENSAM, Paris, Juin 1996.
- Rubin, J. (1994). *Handbook of Usability Testing. How to Plan, Design and Conduct Effective Tests*. New York, John Wiley & Sons.
- Russo, J.E. & Rosen, L.D. (1975). An eye fixation analysis of multialternative choice. *Memory & Cognition*, 3, 267-276.
- Samoylenko, E., McAdams, S. & Nosulenko, V. (1996). Systematic analysis of verbalizations produced in comparing musical timbres. *International Journal of Psychology*, 31(6), 255-278.
- Savoyant, A. (1979). Eléments d'un cadre d'analyse de l'activité : quelques conceptions essentielles de la psychologie soviétique. *Cahiers de Psychologie*, 22, 17-28.
- Scane, R. (1987). A historical perspective. In M.M. Gardiner & B. Christie (Eds.), *Applying Cognitive Psychology to User-Interface Design*. Chichester, UK, Wiley, pp. 57-82.
- Scapin, D.L. & Bastien, J.M.C. (1997). Ergonomic criteria for evaluating the ergonomic quality of interactive systems. *Behaviour & Information Technology*, 16 (4/5), 220-231.
- Sénach, B. (1990). *Evaluation ergonomique des interfaces homme-machine : une revue de la littérature* (Rap. n°1180). Rocquencourt, France, INRIA, mars 1990.

- Shneiderman, B. (1997). *Designing information-abundant websites: issues and recommendations*. College Park, Department of Computer Science & Institute for Systems Research, University of Maryland, 26 February 1997. Disponible depuis: <http://www.cs.umd.edu/projects/hcil/members/bschneiderman/ijhcs/ijhcs.html>.
- Shneiderman, B. (1998). *Designing the User Interface. Strategies for Effective Human-Computer Interaction* (3rd ed.). Reading, MA, Addison Wesley.
- Smith, S.L. & Mosier, J.N. (1986). *Guidelines for designing user interface software* (Rep. N° ESD-TR-86-278). Bedford, MA, Mitre Corporation, August 1986.
- Spérandio, J.C. (1985). L'ergonomie de l'informatique, évolution passée et perspective de recherche. In R. Patesson (Ed.) *Ergonomie et conditions de travail en informatique. Colloque L'ordinateur, l'homme et l'organisation*, Nivelles, Belgique, 1984. Bruxelles, Editions de l'Université de Bruxelles, 17-25.
- Spiller, P. & Lohse, G.L. (1998). A classification of Internet retail stores. *International Journal of Electronic Commerce*, 2(2), 29-56.
- Spool, J.M., Scanlon, T., Schroeder, W., Snyder, C. & DeAngelo, T. (1999). *Web usability. A designer's guide*. San Francisco, CA, Morgan Kaufmann Publishers.
- SSHA (1998). *Evaluation sensorielle. Manuel méthodologique* (2^{ième} ed.). Paris, Lavoisier, coll. Tec et Doc.
- Stanton, N. & Baber, C. (1996a). Factors affecting the selection of methods and techniques prior to conducting a usability evaluation. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 39-48.
- Stanton, N. & Baber, C. (1996b). Task analysis for error identification: applying HEI to product design and evaluation. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 215-223.
- Stewart, D.W. & Shamdasani, P.N. (1990). *Focus Groups. Theory and Practice*, Applied Social Research Methods Series, vol. 20. Newbury Park, CA, Sage Publications.

- Stone, H., Sidel, J., Oliver, S., Woolsey, A. & Singleton, R.C. (1974). Sensory evaluation by quantitative descriptive analysis. *Food Technology*, 28(11), 24-34.
- Sullivan, P. (1991). Multiple methods and the usability of interface prototypes : the complementarity of laboratory observation and focus groups. *Proceedings of SIGDOC'91, the 9th Annual International Conference on Systems Documentation*, Chicago, IL, 10-12 October 1991. New York, ACM Press, 106-112.
- Sutcliff, A. (2000). Pay attention ! Or how to make the users find the message in multimedia. In B.S. Scapin & E. Vergison (Eds.), *Actes de la Conférence ERGO-IHM 2000*, Biarritz, France, 3-6 octobre 2000. Bidart, France, CRT ILS & ESTIA, 2-12.
- Teo, H.-H., Oh, L.-B., Liu, C. & Wei, K.-K. (2003). An empirical study of the effects of interactivity on web user attitude. *International Journal Human-Computer Studies*, 58(3), 281-305.
- Teo, T.S.H. (2002). Attitudes toward online shopping and the Internet. *Behaviour & Information Technology*, 21(4), 259-271.
- Thomas, B. (1996). 'Quick and dirty' usability tests. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability Evaluation in Industry*. London, Taylor & Francis, pp. 107-114.
- Tilson, R., Dong, J., Martin, S. & Kieke, E. (1998). Factors and principles affecting the usability of four e-commerce sites. *Virtual Proceedings of the 4th Conference on Human Factors and the Web*, 5 June 1998. Basking Ridge, NJ, AT&T Labs. Disponible depuis: <http://www.research.microsoft.com/users/marycz/hfweb98/tilson/>.
- Urban, G.L. & Hauser, J.R. (1993). *Design and Marketing of New Products* (2nd ed.). Englewood Cliffs, NJ, Prentice-Hall.
- Urdapilleta, I. (2001a). Introduction générale. In I. Urdapilleta, C. Ton Nu, C. Saint Denis & F. Huon de Kermadec (Eds.), *Traité d'évaluation sensorielle. Aspects cognitifs et métrologiques des perceptions*. Paris, Dunod, pp. XV-XXVIII.
- Urdapilleta, I. (2001b). Mise au point d'une liste de descripteurs par décision de groupe. In I. Urdapilleta, C. Ton Nu, C. Saint Denis & F. Huon de Kermadec (Eds.), *Traité d'évaluation sensorielle. Aspects cognitifs et métrologiques des perceptions*. Paris, Dunod, pp. 91-106.

- Urdapilleta, I. (2001c). Catégorisation et cotation, les échelles hiérarchiques à items verbaux. In I. Urdapilleta, C. Ton Nu, C. Saint Denis & F. Huon de Kermadec (Eds.), *Traité d'évaluation sensorielle. Aspects cognitifs et métrologiques des perceptions*. Paris, Dunod, pp. 107-124.
- Urdapilleta, I. & Bouard, C. (2001). Analyse de profils sensoriels par une méthode basée sur le principe de diffusion et la théorie des ensembles flous. In I. Urdapilleta, C. Ton Nu, C. Saint Denis & F. Huon de Kermadec (Eds.), *Traité d'évaluation sensorielle. Aspects cognitifs et métrologiques des perceptions*. Paris, Dunod, pp. 125-146.
- Urvanstev, L.P. (1979). Eksperimentalnoe issledovanie protsessov vospriyatia i interpretatsii rentgenogramm (Recherche expérimentale des processus de perception et d'interprétation des radiographies), *Psikhologicheskie problemy ratsionalizatsii deyatel'nosti, svyazannoj s opoznaniem obrazov*, (Les problèmes psychologiques de rationalisation de l'activité de reconnaissance des images). Yaroslavl, Université de Yaroslavl
- Urvanstev, L.P. (1994). Obraznye komponenty v deyatel'nosti rentgenologa (Les composantes de l'image dans l'activité du radiologue). *Psikhitcheskij obraz: stroenie, mekhanizmy, funkcionirovanie i razvitie*. (Image psychique: conception, mécanismes, fonctionnement et développement. Actes de la deuxième conférence internationale de B.F. Lomov, vol. 2, 105-107). Moscou, Institut de Psychologie de l'Académie des Sciences de Russie.
- Van Vianen, E., Thomas, B. & Van Nieuwkastele, M. (1996). A combined effort in the standardization of user interface testing. In P.W. Jordan, B. Thomas, B.A. Weerdmeester & I.L. McClelland (Eds.), *Usability in Industry Evaluation*. London, Taylor & Francis, pp. 7-17.
- Vanderdonckt, J. (1995). *Guide ergonomique des interfaces homme-machine*. Namur, Belgique, Presses Universitaires de Namur.
- Vergnaud, G. (1999). On n'a jamais fini de relire Vygotski et Piaget. In Y. Clot (Ed.), *Avec Vygotski*. Paris, La Dispute/SNEDIT, pp. 45-58.
- Vermersch, P. (1994). *L'entretien d'explicitation en formation initiale et en formation continue*. Paris, ESF éditeur, coll. Pédagogies.

- Vermersch, P. (2000). *L'entretien d'explicitation* (3^{ième} ed.). Issy-les-Moulineaux (Hauts-de-Seine), France, ESF éditeur, coll. Pédagogies Outils.
- W3C (1999). Web content accessibility guidelines 1.0. W3C recommandation 5 May 1999, sur le site Internet *W3C The World Wide Web Consortium*, disponible depuis: <http://www.w3.org/TR/WAI-WEBCONTENT/>.
- Warfield, J.N. (1976). *Societal Systems : Planning, Policy and Complexity*. New York, Wiley.
- Weill-Fassina, A., Rabardel, P. & Dubois, D. (1993). *Représentation pour l'action*. Toulouse, Octarès.
- Whitefield, A., Wilson, F. & Dowell, J. (1991). A framework for human factors evaluation. *Behaviour & Information Technology*, 10(1), 65-79.
- Wight, D. (1994). Boys' thoughts and talk about sex in a working class locality of Glasgow. *Sociological Review*, 42, 702-737.
- Williams, A.A. & Langron, S.P. (1984). The use of free-choice profiling for the evaluation of commercial ports. *Journal of the Science of Food and Agriculture*, 35, 558-568.
- Woolrych, A. & Cockton, G. (2001). Why and when five test users aren't enough. In J. Vanderdonckt, A. Blandford & A. Derycke (Eds.), *Interaction sans frontières. Actes de IHM-HCI 2001: Conférence conjointe AFIHM-BCS sur l'Interaction Homme-Machine*, Lille, France, 10-14 septembre 2001. Toulouse, France, Cépaduès Editions, 105-108.
- Wright, P.C. & Monk, A.F. (1991). The use of think-aloud evaluation methods in design. *SIGCHI Bulletin*, 23(1), 55-57.
- Yoo, B. & Kim, J. (2000). Experiment on the effectiveness of link structure for convenient cybershopping. *Journal of Organizational Computing and Electronic Commerce*, 10(4), 241-256.

ANNEXES

ANNEXES DU CHAPITRE 3 :

GUIDE DE L'ENTRETIEN INDIVIDUEL

Accueil :

Explication du déroulement de l'entretien. Seconde demande d'assentiment du sujet par rapport au fait d'être enregistré « audio-visuellement » (première demande lors du recrutement).

Entretien préalable sur les caractéristiques du sujet (fréquence d'utilisation d'Internet, etc.) et remplissage du formulaire s'y rapportant par l'intervieweur.

Phase de verbalisation sans utilisation du produit/service

1. Technique d'élicitation associative :

« Ecrivez tout ce qui vous vient à l'esprit, sans rien censurer, concernant les bons et les mauvais aspects d'un site d'achat en ligne ? »

Le sujet dispose de cinq minutes pour réaliser cela. Un court entretien de clarification de chaque élément écrit est ensuite mené.

2. Entretien d'explicitation :

« Vous souvenez-vous de votre dernière utilisation d'un site d'achat en ligne ? Si oui, je vous propose de re-dérouler point par point cette expérience d'utilisation. Quelle est la toute première opération que vous ayez réalisée ? Comment êtes-vous arrivés sur le site ? par un moteur de recherche, le site faisait partie de vos favoris ?... »

Si le sujet ne se souvient pas de sa dernière expérience ou préfère en évoquer une autre, l'expérimentateur réalise l'entretien d'explicitation à partir de cette autre expérience.

3. Phase de synthèse écrite (fiches-critère) :

« Nous allons reprendre une par une chacune des idées que vous avez évoquées et je vous demande de les reformuler sous la forme de critères avec leur intitulé et leur définition. »

- Description de l'exemple du téléphone mobile :

« Les différents critères de jugement du téléphone mobile pourrait, par exemple, être :

- Coût : prix d'achat, tarifs de l'abonnement, tarifs des communications, prix des accessoires, ...
- Esthétique : couleur, forme, mélodies des sonneries, ...
- Poids : poids qui soit le moins lourd possible pour ne pas alourdir son sac, et pour qu'il ne soit pas lourd dans la main, quantité en grammes.

De la même façon nous vous invitons à décrire vos critères de jugement d'un site d'achat en ligne. Commencez, par exemple, par remplir des éléments de définition puis vous trouverez un titre à ce critère, »

Le sujet est accompagné dans cette phase d'écriture. L'intervieweur relit les notes qu'il a prises durant l'entretien d'explicitation pour permettre au sujet de se remémorer tout son récit d'expérience vécue.

Phase de verbalisation avec utilisation du produit/service

4. Méthode des verbalisations libres :

« Nous vous proposons maintenant d'abord d'explorer la page d'accueil, puis de chercher un produit de votre choix sur un site d'achat en ligne puis sur un autre, que vous allez ainsi tous deux comparer. »

« Vous pouvez choisir un thème parmi les quatre thèmes suivants :

- Culturel (Amazone, Fnac),
- Mode (la Redoute, 3 Suisses),
- Voyages (Degriptour, Travelprice),
- Vins (Cave-net, Vins de France) ».

- Tâche d'exploration de la page d'accueil :

Site 1 : « Pendant que vous explorez la page d'accueil du site d'achat en ligne, pouvez-vous signaler tout haut, votre impression générale du site, vos remarques, vos appréciations. »

Des relances neutres sont prévues pour relancer les verbalisations, par exemple, « Pouvez-vous dire tout haut vos impressions ».

Site 2 : « *Nous vous demandons maintenant de réaliser la même chose sur l'autre site.* »

A la fin de la tâche : « *Pouvez-vous décrire les différences et ressemblances entre les deux sites.* »

- Tâche de recherche d'un produit :

« *Choisissez maintenant un produit que vous allez rechercher sur l'un puis sur l'autre site, ainsi que de remplir les bons de commandes sans pour autant procéder à l'achat.* »

Site 1 : « *En parcourant l'ensemble des informations, pouvez-vous dire ce qui vous plaît et/ou déplaît, ce que vous pensez des fonctionnalités proposées.* »

Des relances neutres sont prévues pour relancer les verbalisations, par exemple, « *Pouvez-vous dire tout haut vos impressions* ».

Site 2 : « *Nous vous demandons maintenant de réaliser la même chose sur l'autre site.* »

A la fin de la tâche : « *Pouvez-vous décrire les différences et ressemblances entre les deux sites. Quel est le site qui répond le plus à vos critères et pourquoi ?* »

5. Phase de synthèse écrite (seconde liste de fiches-critère) :

« *Nous allons maintenant reprendre l'ensemble des idées que vous avez évoquées et je vous demande, à nouveau, de les reformuler sous la forme de critères avec leur intitulé et leur définition.* »

« *N'hésitez pas, si vous le souhaitez, à reformuler deux critères identiques avec un nouvel intitulé et une nouvelle définition* ».

Le sujet est accompagné dans cette phase d'écriture. L'intervieweur relit les notes qu'il a prises durant la séquence d'utilisation/verbalisations libres pour permettre au sujet de se remémorer les éléments parcourus sur le site marchand et ses commentaires.

GUIDE DU FOCUS GROUP

Accueil :

Explication du déroulement de l'entretien. Rappel par rapport au fait qu'ils sont enregistrés (première demande d'assentiment concernant les enregistrements audio-visuels lors du recrutement).

Remplissage d'un formulaire sur les caractéristiques de chaque participant (fréquence d'utilisation d'Internet, etc.) dès leur arrivée dans la salle d'attente.

Phase de verbalisation sans utilisation du produit/service

1. (et 3.) Technique d'élicitation associative et création et nomination de critères collectifs :

« Ecrivez, chacun, tout ce qui vous vient à l'esprit, sans rien censurer, concernant les bons et les mauvais aspects d'un site d'achat en ligne ? »

Au bout de cinq minutes environ, les participants sont invités, un par un, à énoncer les éléments qu'ils ont écrits. Tous les éléments sont inscrits et numérotés par l'animateur sur le « paper-board ».

Puis, travail de catégorisation : *« Nous allons maintenant prendre les éléments écrits, un par un, et les regrouper selon leur similitude. Nous commençons par les cinq premiers, est-ce que certains de ces éléments vous semblent exprimer la même idée ?... »*

Inscription des numéros des éléments sur des fiches cartonnées qui sont placées sur un grand tableau « autocollant » (à côté du « paper-board »).

« Nous poursuivons avec les cinq suivants. Le premier élément exprime-t-il une idée similaire à un des éléments déjà placés sur le tableau, ... ».

Puis discussion modérée par l'animateur qui cherche à faire ressortir les points de convergence dans les interactions pour amener le groupe à un consensus autour de regroupements. Par ailleurs, il reformule les points de divergence pour les clarifier et propose de laisser un élément de côté si le groupe ne parvient pas à un consensus le concernant.

Il invite également à donner un titre aux regroupements formés au fur et à mesure, même si ces regroupements évoluent avec l'introduction d'autres éléments (scission en deux d'un regroupement, combinaison, d'un regroupement avec un autre). Ces titres équivalent aux fiches-critère de l'entretien individuel.

2. Discussions à partir d'expériences vécues :

Le précédent travail de catégorisation doit aussi générer des discussions de confrontation où les participants vont s'expliquer leur point de vue à l'aide d'expériences personnelles. L'animateur reformule le récit de ces expériences si nécessaire et pour les encourager.

3. L'étape 3 :

Cf. étape 1 de la méthodologie du *focus group*. Nomination des regroupements ou des critères de jugement collectifs, équivalant à l'écriture de fiches-critères de l'entretien individuel.

Phase de verbalisation avec utilisation du produit/service

4. La méthode des verbalisations libres :

« Nous vous proposons maintenant d'abord d'explorer la page d'accueil, puis de chercher un produit de votre choix sur un site d'achat en ligne puis sur un autre, que vous allez ainsi tous deux comparer. »

« Vous pouvez choisir un thème parmi les quatre thèmes suivants :

- Culturel (Amazon, Fnac),*
- Mode (la Redoute, 3 Suisses),*
- Voyages (Degriptour, Travelprice),*
- Vins (Cave-net, Vins de France) ».*

« Il faut un volontaire pour naviguer sur les sites marchands à partir de ce micro-ordinateur. Les pages Internet sont projetées sur l'écran et pourront donc être observés par les autres. »

- Tâche d'exploration de la page d'accueil :

Site 1 : *« Pendant que Monsieur ou Madame navigue sur la page d'accueil du site d'achat en ligne, pouvez-vous commenter tout haut, votre impression générale du site, vos remarques, vos appréciations. ».*

Site 2 : « *Nous vous demandons maintenant de réaliser la même chose sur l'autre site.* »

A la fin de la tâche : « *Pouvez-vous décrire les différences et ressemblances entre les deux sites.* »

Modération des discussions spontanées occasionnées par la visualisation des pages Internet.

- Tâche de recherche d'un produit :

« *Choisissez maintenant un produit que vous allez rechercher sur l'un puis sur l'autre site.* »

Site 1 : « *Pouvez-vous dire ce qui vous plaît et/ou déplaît, ce que vous pensez des fonctionnalités proposées.* »

Site 2 : « *Nous vous demandons maintenant de réaliser la même chose sur l'autre site.* »

A la fin de la tâche : « *Pouvez-vous décrire les différences et ressemblances entre les deux sites.* »

Modération des discussions spontanées occasionnées par la visualisation des pages Internet.

5. Discussion de synthèse sur les critères de jugements :

« *Nous vous proposons de faire le point sur vos critères de jugement suite à cette séquence d'utilisation/observation des sites marchands.* »

L'animateur recherche les consensus dans le groupe et dresse une liste de critères sur le « paper-board », équivalant à la seconde liste de fiches-critères de l'entretien individuel.

LISTE EXHAUSTIVE DE DIMENSIONS UTILISATEURS ELABOREES A PARTIR DES DONNEES DES ENTRETIENS INDIVIDUELS ET DES FOCUS GROUPS

équivalent à la grille d'analyse des verbalisations des entretiens individuels et des focus groups

Dimensions d'évaluation liées plus particulièrement aux sites Internet marchands :

N° de dimension	Intitulé de la dimension
1	Notoriété
2	Facilité d'accès
3	Clarté du service rendu et informations qui rassurent dès la page d'accueil
4	Accès et qualité du contenu des informations de vente et de sécurité (avant la commande)
5	Existence et qualité de la relation humaine au-delà de la machine (conseiller en vente)
6	Diversité et qualité des produits proposés
7	Prix des produits proposés
8	Différences d'Internet / commerce classique (même choix de prix, produits, promotions)
9	Propositions de promotions, de nouveautés et de divertissements
10	Qualité de la description et des informations liées aux produits (caractéristiques, photo, disponibilité, prix, etc.)
11	Existence de différentes fonctions de recherche (thématique de départ)
12	Facilité de la recherche de produits (saisie et cheminement)
13	Pertinence et facilité de lecture des résultats de la recherche
14	Informations supplémentaires et propositions d'autres produits
15	Facilité de la sélection des produits à la commande (panier)
16	Facilité et rapidité de la commande (inscription/commande/paiement)
17	Informations simultanées sur la sécurité/vente et feed-back sur la commande
18	Sécurité et modes de paiement
19	Confidentialité des données personnelles (vie privée)
20	Informations personnalisées pour le client et enregistrement (identifiant)
21	Possibilité de pré-commande, suivi de la commande
22	Efficacité dans l'expédition (disponibilité – livraison)
23	Efficacité des services d'après vente (du livreur, des retours de produits, etc.)

Dimensions d'évaluation génériques, relatives à tout site Internet :

24	Aide
25	Clarté du vocabulaire
26	Rapidité et facilité de navigation
27	Qualité (pertinence, mise en évidence, lisibilité) des informations
28	Esthétique du site
29	Publicités
30	Mise à jour du site
31	Compatibilité avec les logiciels courants

RESSOURCES HUMAINES NECESSAIRES ET COUTS DE REALISATION DE DIFFERENTES ETAPES DE L'ENTRETIEN INDIVIDUEL ET DU *FOCUS GROUP*

Nous explicitons ici les ressources humaines et les coûts engagés à chaque étape de réalisation de l'entretien individuel et du *focus group* (voir les récapitulatifs dans les tableaux 100 à 103, page suivante).

- La préparation :

Dans le temps de préparation, sont comptés trois jours de documentation (le professionnel de l'évaluation rassemble divers documents en interne dans l'entreprise ou des articles de la presse spécialisée, etc.) et deux jours de préparation du guide d'entretien, des critères de recrutement des participants, etc. Globalement, les mêmes durées de préparation sont nécessaires que l'étude se réalise par entretiens individuels ou *focus groups* (un bon guide de *focus group* prend autant de temps à construire qu'un bon guide d'entretien individuel, les caractéristiques des participants sont les mêmes dans les deux cas, etc.).

- Le recrutement :

Il faut deux jours environ au « recruteur » (personne du service d'évaluation formée à cette activité et ayant développé une importante base d'utilisateurs) pour recruter les huit participants, qu'il s'agisse de les réunir lors d'une même soirée ou de trouver une date de rendez-vous individuel. Généralement, il est plus difficile de réunir plusieurs participants pour un même créneau horaire (Crabtree *et al.*, 1993), mais dans le cas d'études concernant des produits grand public, il s'agit de personnes « lambdas » et non d'individus très ciblés et/ou très pris dans leur emploi du temps. Avec une base de participants potentiels suffisamment large (ce qui est le cas), ces huit personnes sont assez facilement rassemblées.

- La passation des entretiens :

Les durées de passation « brutes » sont de 12 heures pour les huit entretiens individuels et de deux heures et demie pour un *focus group*. Ces durées sont ramenées en termes de demi-journée en ajoutant les temps morts entre chaque entretien ce qui est plus réaliste concernant la location d'une salle et la mobilisation de l'intervieweur/animateur (professionnel de l'évaluation) et de son assistant. Le temps de l'installation (un jour pour une série de huit entretiens individuels ou la journée pour le soir du *focus group*, la soirée équivalent à une demi-journée) est également ajouté.

- Le traitement des données (et rédaction du rapport) :

Nous avons appris des professionnels de France Télécom comme de ceux de cabinets d'évaluation que le temps de traitement des données correspond à environ deux fois le temps de passation pour un traitement des données assez fin pour une pratique « de terrain » (à partir des prises de notes et de la visualisation des vidéos) pour les entretiens individuels et à environ trois fois le temps de passation pour le *focus group* pour le même type de traitement. A tout cela, est ajouté deux jours de rédaction du rapport de synthèse, pour les deux techniques.

- Les coûts financiers :

Enfin, nous utilisons la grille de salaires de France Télécom, les prix de location des salles pratiqués par le service d'évaluation (quand ils les louent à d'autres services), les prix des repas distribués lors des *focus groups* et le montant des chèques cadeaux remis à tous les participants. Auparavant, nous avons vérifié que ces frais étaient similaires ou proportionnels à ceux pratiqués par des cabinets d'évaluation parisiens.

Les différents temps, ressources humaines (en hommes-heures, HH) et coûts sont récapitulés et détaillés dans les tableaux 100 à 103 ci-dessous :

Tableau 100 : Détail des ressources humaines (en hommes-heures) pour les différentes étapes de réalisation de l'entretien individuel

Différentes étapes	Préparation	Recrutement	Passation	Installation, animation	Traitement + rapport
Détail	3 jours de documentation, 2 jours de préparation, 1 évaluateur	2 jours, 1 « recruteur »	2½ jours pour faire les 8EI, 1 participant	1 j d'installation avec 1 technicien, 2½ j d'animation avec 1 évaluateur et 1 technicien	2½ jours × 2 de traitement, 2 jours de rédaction, 1 évaluateur
Détail nb d'hommes × heures (HH)	5j × 8h × 1HH	2j × 8h × 1HH	2½j × 8h × 1HH	1j × 8h × 1HH + 2½j × 8h × 2HH	7j × 8h × 1HH
Total (en HH)	40 HH	16 HH	20HH	48HH	56HH

Légende : EI = Entretien individuel, HH = hommes-heures.

Tableau 101 : Détail des ressources humaines (en hommes-heures, HH) pour les différentes étapes de réalisation du focus group

Différentes étapes	Préparation	Recrutement	Passation	Installation, animation	Traitement + rapport
Détail	3 jours de documentation, 2 jours de préparation, 1 évaluateur	2 jours, 1 « recruteur »	½ jour pour faire le FG, 8 participants	1 j d'installation avec 1 technicien, ½ j d'animation avec 1 évaluateur, 1 assistant et 1 technicien	½ jour × 3 de traitement, 2 jours de rédaction, 1 évaluateur
Détail nb d'hommes × heures (HH)	5j × 8h × 1HH	2j × 8h × 1HH	½j × 8h × 8HH	1j × 8h × 1HH + ½j × 8h × 3HH	3½j × 8h × 1HH
Total (en HH)	40 HH	16HH	32HH	20HH	28HH

Légende : FG = focus group, HH = hommes-heures.

Tableau 102 : Détail des coûts financiers de réalisation des différentes étapes de l'entretien individuel

Différentes étapes	Préparation	Animation	Traitement+ rapport	Recrutement, Indemnisations	Matériel
Détail	3 jours de documentation 2 jours de préparation, 1 évaluateur (702 €j)	2½ jours pour animer les 8EI, 1 évaluateur (702 €j)	2½ jours × 2 de traitement, 2 jours de rédaction, 1 évaluateur (702 €j)	2 jours, 1 « recruteur » (560 €j), 8 chèques cadeaux (à 23 €)	3½ jours (1j installation + 2½ j passation) de location d'une petite salle (800 €) comprenant les services du technicien
Détail calcul (en €)	5j × 702 €	2½ × 702 €	7j × 702 €	2j × 560 € + 8 × 23 €	3½ j × 800 €
Total (en €)	3510 €	1755 €	4914 €	1304 €	2800 €

Tableau 103 : Détail des coûts financiers de réalisation des différentes étapes du focus group

Différentes étapes	Préparation	Animation	Traitement + rapport	Recrutement, Indemnisations	Matériel
Détail	3 jours de documentation et 2 jours de préparation 1 évaluateur (702 €j),	½ jours pour animer le FG 1 évaluateur (702 €j) et 1 assistant (702 €), 3 repas (avec celui technicien à 12 €)	½ jour × 3 de traitement, 2 jours de rédaction, 1 évaluateur (702 €j)	2 jours, 1 « recruteur » (560 €j), 8 chèques cadeaux (à 38 €), 8 repas (12 €)	1 jour (la journée d'installation pour le soir) de location d'une grande salle (900 €) comprenant les services du technicien
Détail calcul (en €)	5j × 702 €	0,5j × 2 × 702 € + 3 × 12 €	3½j × 702 €	2j × 560 € + 8 × 38 € + 8 × 23 €	1j × 900 €
Total (en €)	3510 €	738 €	2457€	1608 €	900€

ANNEXES DU CHAPITRE 4 :

ENTRETIEN AVEC UN INGENIEUR CONCEPTEUR DE FTA COMMENTANT LE RAPPORT D'EVALUATION DE LA SATISFACTION DES UTILISATEURS D'UNE APPLICATION LOGICIELLE POUR L'ACCES A INTERNET :

L'étude d'évaluation était commandée par la direction commerciale d'une branche de France Télécom ayant conçu la nouvelle version d'une application logicielle pour l'accès à Internet. Cette étude s'adresse donc à des commerciaux mais a également été transmise aux concepteurs du produit qui ont réalisé les développements informatiques de son processus d'installation et de l'interface de départ avec boutons d'accès aux principales fonctionnalités.

Un entretien semi directif a été mené avec un de ces concepteurs qui a commenté le rapport (non publiable pour des raisons de confidentialité) de cette étude en précisant ce qui était utile pour poursuivre l'amélioration de la conception et ce qui ne l'était pas.

Résumé des besoins en connaissances des concepteurs révélés lors de cet entretien :

- L'information sur le détail de l'échantillon des utilisateurs et de la méthodologie d'étude est inutile pour les concepteurs (éventuelle lecture à titre indicatif),
- Les données sur les usages ou la satisfaction par profils utilisateur (ex. les 13-25 ans) ne peuvent être utilisées que si elle révèle un problème critique (ex. 70% de non-utilisation d'une fonction). Sinon, il s'agit d'une information à titre indicatif qui ne peut donner lieu à une prise de décisions précises quant à la transformation du produit,
- Les concepteurs ont besoin :
 - de données précises sur les difficultés rencontrées,
 - d'éléments chiffrés pour évaluer l'importance relative de chaque problème et prendre des décisions en termes de transformation. Par exemple, les verbatim ne constituent pas des résultats qui soient véritablement utilisables. Le nombre d'utilisateurs ayant rencontré une difficulté est en revanche très utile,
- L'évaluation selon divers critères (dimensions d'évaluation) leur est utile, surtout si elle est accompagnée des explications des causes (positives ou négatives) pour identifier les points à améliorer ou à conserver,

- La carte ou « mapping » situant les points forts et points faibles du produit est très utile.

Description sommaire de l'étude d'évaluation commentée

L'étude d'évaluation a tenu compte des principales orientations de conception de départ sans faire une étude approfondie a priori des besoins en connaissances des concepteurs ou des commerciaux. Par exemple, une des orientations de conception de départ était d'améliorer la facilité d'utilisation du produit pour que les utilisateurs fassent moins appel à la « hot line ». L'évaluation a donc porté, entre autres, sur les difficultés rencontrées lors de l'installation du produit en précisant le taux d'utilisateurs ayant appelé la « hot line ».

Il s'agissait d'une étude quantitative auprès d'un très large panel d'utilisateurs réels du produit (l'ayant reçu depuis un certain temps), selon divers critères (non précisés pour des raisons de confidentialité) mis sur la forme d'échelle de Likert en 5 points et à items verbaux (ex. très satisfaisant, assez satisfaisant, ne se prononcent pas, peu satisfaisant, pas du tout satisfaisant). L'évaluation de l'importance relative de chacun des critères était également prévue. Enfin, quelques entretiens téléphoniques avec certains de ces utilisateurs ont été menés.

Les critères eux-mêmes ont été révélés grâce à une étude amont auprès d'utilisateurs et tenaient compte des besoins (très généraux) en connaissances des concepteurs (en rapport avec les orientations principales de conception). Des explications détaillées de ces critères, l'identification et explication de certaines difficultés rencontrées et des types de profils d'utilisateurs (ex. ayant tel usage, etc.) accompagnaient ces résultats.

L'évaluation portait sur trois grands éléments du produit :

- sa présentation multimédia (une démonstration des fonctionnalités du produit),
- son processus d'installation,
- l'interface (de départ) avec boutons d'accès aux principales fonctionnalités.

Commentaires du concepteur quant aux résultats d'évaluation

Voici une description sommaire des commentaires du concepteur concernant les différents résultats de l'étude d'évaluation.

Commentaires par rapport aux résultats concernant la présentation multimédia

Les résultats de l'étude concernent les :

- 1/ Taux de satisfaction quant à la présentation,
- 2/ Profil d'utilisateur ayant le plus consulté la présentation,
- 3/ Taux d'utilisateurs ayant consulté la présentation avant de lancer l'installation,
- 4/ Extraits de verbatim concernant la présentation.

- 1/ Taux de satisfaction quant à la présentation : 95 %

Les concepteurs ont besoin que cette donnée soit corrélée avec des éléments détaillés d'explication des raisons de la satisfaction, dans ce cas pour pouvoir conserver (voire accentuer) les points forts de la présentation

- 2/ Profil d'utilisateur ayant le plus consulté la présentation : les femmes

Cette donnée est utile seulement à titre indicatif. Cela ne donne pas d'indications pour entreprendre une transformation.

- 3/ Taux d'utilisateurs ayant consulté la présentation avant de lancer l'installation : 2/5 des utilisateurs

La majorité des utilisateurs lancent donc l'installation et utilisent l'interface de départ sans rien en savoir. Les concepteurs avaient, en fait, anticipé cette donnée, c'est-à-dire qu'ils avaient prévu un produit d'une très grande simplicité d'utilisation, le plus didactique possible. Leur orientation de conception de départ était de concevoir pour des néophytes. Cette donnée vient donc confirmer que l'orientation de conception choisie est en accord avec l'usage effectif des utilisateurs.

- 4/ Extraits de verbatims concernant la présentation

Exemple : « *il faudrait peut-être des exemples interactifs qui imagent facilement les informations* ». Problème : il n'y a pas de données sur le poids relatif de chaque verbatim. Les

concepteurs ne savent pas quelle importance leur accorder et ne peuvent donc pas en tirer de conclusion sur les actions de transformation à mener.

La conclusion de l'étude est qu'il manque quelque chose pour préparer le néophyte. Cependant une autre contrainte s'impose au niveau de la conception : techniquement, rajouter des exemples interactifs dans la présentation multimédia n'est pas une bonne solution.

Commentaires par rapport aux résultats concernant le processus d'installation et l'aide

Orientation de conception de départ : efficacité de l'installation et notamment de l'aide qui s'y rapporte pour éviter les recours à la « hot line ».

Les résultats de l'étude concernent :

- 1/ Graphiques de proportion (« en fromages ») sur le nombre d'utilisateurs ayant eu des difficultés lors de l'installation et solutions utilisées pour résoudre la difficulté,
- 2/ Graphiques de proportion (« en fromages ») sur l'utilisation de l'aide,
- 3/ Graphiques de proportion (« en fromages ») sur la lecture des écrans,
- 4/ Liste détaillée des difficultés rencontrées avec le décompte du nombre d'utilisateurs l'ayant rencontrée,
- 5/ Profil des utilisateurs ayant (ou non) utilisé l'aide.

• 1/ Les graphiques de proportion (« en fromages ») concernant les difficultés lors de l'installation

Les graphiques révèlent le :

- nombre d'utilisateurs ayant trouvé la solution seul,
- nombre d'utilisateurs ayant eu recours à l'aide,
- nombre d'utilisateurs ayant eu recours à la « hot line »,
- nombre d'utilisateurs ayant eu recours à d'autres sources d'informations.

Selon le concepteur, il manque une corrélation avec des résultats concernant l'efficacité de l'aide. Il manque le nombre d'utilisateurs ayant eu recours d'abord à l'aide (ou à un autre recours, ex. un proche) et ayant ensuite eu recours à la « hot line » car n'ayant pas trouvé l'information nécessaire. Il faudrait également des données sur le détail des raisons pour lesquelles il n'ont pas trouvé cette information dans l'aide.

- 2/ Les graphiques de proportion (« en fromages ») sur l'utilisation de l'aide

Les graphiques ne concernent que :

- les utilisateurs ayant utilisé l'aide,
- les utilisateurs ayant utilisé l'aide à toutes les étapes,
- n'ayant pas utilisé l'aide.

Au total, seulement la moitié des utilisateurs ont utilisé l'aide. Cette donnée est très importante pour les concepteurs car une autre orientation pour la conception de la nouvelle version du produit était l'efficacité de l'aide, d'où un travail important sur cette aide (vocabulaire, enchaînement des écrans, etc.).

Finalement, le concepteur ne voit pas d'autre solution que « d'éduquer » l'utilisateur à utiliser davantage l'aide (probablement en raison du travail important fourni sur la conception de l'aide). Cependant, le concepteur lui-même reconnaît qu'il n'a pas non plus le réflexe d'utiliser l'aide et doute donc de son précédent propos.

Il manque aussi des données concernant les détails des difficultés d'utilisation de l'aide.

- 3/ Les graphiques de proportion (« en fromages »)

Les graphiques révèlent le :

- nombre d'utilisateurs ayant lu chacun des écrans (de l'installation),
- nombre d'utilisateurs n'ayant lu aucun écran.

La moitié des utilisateurs ne lisent pas les écrans. Cela signifie que la plupart des utilisateurs cliquent sur « OK » sur chaque écran et ne savent pas selon quels paramètres ils installent le produit.

Les concepteurs avaient prévu que par défaut le mode « néophyte » soit installé. Cette donnée vient donc confirmer que les concepteurs ont fait le bon choix de conception en rapport avec l'usage des utilisateurs.

- 4/ Liste détaillée des difficultés

Ces données sont très utiles, en particulier le nombre d'utilisateurs ayant rencontré chaque difficulté. Les éléments chiffrés permettent d'estimer l'importance de chaque difficulté et sont indispensables pour les concepteurs. Par contre, ils ne retiennent que les difficultés sur lesquelles ils peuvent techniquement agir.

- 5/ Profil utilisateur

Les jeunes entre 13 et 25 ans ont le moins utilisé l'aide (environ 70%). Il s'agit d'une donnée retenue par les concepteurs, car ils doivent donc essayer d'agir pour que cette population aient davantage envie d'utiliser l'aide.

Commentaires par rapport aux résultats concernant l'interface avec boutons d'accès aux principales fonctionnalités

Les résultats de l'étude concernent :

- 1/ la satisfaction des utilisateurs quant au concept,
- 2/ l'appréciation des icônes selon divers critères.

- 1/ La satisfaction des utilisateurs quant au concept

L'évaluation de la satisfaction est utile pour les concepteurs.

Pour la plupart, les utilisateurs étaient très satisfaits du concept, à savoir le regroupement de toutes les grandes fonctionnalités sur une même interface. Cela confirme aux concepteurs qu'ils ont pris la bonne orientation concernant le concept.

- 2/ Résultats d'appréciation des icônes selon divers critères

L'évaluation fait ressortir que les utilisateurs ont des difficultés dans la compréhension de certaines icônes, en particulier concernant un bouton. Les concepteurs en déduisent qu'il faudrait (peut-être) une meilleure explication de l'utilisation de ces icônes dans la présentation multimédia.

Commentaires par rapport à la conclusion de l'étude d'évaluation

La carte de synthèse ou « mapping » des divers points faibles et points forts du produit en lien avec les trois orientations de conception de départ est très appréciée par le concepteur. Il faudrait cependant améliorer la lisibilité de ce « mapping ».

EXTRAIT DE LA LISTE EXHAUSTIVE DE DIMENSIONS UTILISATEURS DE SITES INTERNET MARCHANDS AVEC LEURS SOUS-DIMENSIONS, AVANT LA TRANSFORMATION DE SURFACE

N° des dimensions	Intitulé de la dimension et numéro et intitulé de ses sous-dimensions
11	Existence de différentes fonctions de recherche (thématique de départ)
	11a. Possibilité de recherche par rubriques
	11b. Possibilité de recherche par « mots clefs »
	11c. Possibilité de recherche par formulaire multi-choix (ville départ, horaires, prix, etc.)
	11d. Possibilité de recherche avancée (combinaison de champs par « mots clefs », ex : titre+auteur) ou guidée (par conseils)
	11e. Possibilité de recherche par référence catalogue
	11f. Possibilité d'accès aux sous-catégories de recherche dès la page d'accueil (ex : pop-ups des sous-menus sur les onglets)
12	Facilité de la recherche de produits (saisie et cheminement)
	12a. Pertinence du choix des rubriques et dans l'affinement proposé
	12b. Efficacité et fonctions du moteur de recherche (recherche par orthographe voisine ou précisément sur l'orthographe donnée, recherche sur tous les mots-clefs de façon indissociable)
	12c. Facilité d'utilisation des menus déroulants et multi-choix ou champs
	12d. Existence de guidage et efficacité du formulaire multi-choix ou du moteur de recherche (ex : propositions de voyages en Afrique du Sud au départ de Paris même si la demande était au départ de Nantes ; propositions d'autres recherches en cas de manque de produits)
	12e. Facilité d'évolution dans la recherche, en général (pas de chemins illogiques)
	12f. Rapidité d'aboutissement de la recherche (3 clics)
13	Pertinence et facilité de lecture des résultats de la recherche
	13a. Existence de présélections et classifications des produits présentés
	13b. Qualité de la présentation des listes de résultats (petite photo, repère par numérotation des pages, etc.)
14	Informations supplémentaires et propositions d'autres produits
	14a. Possibilité d'informations supplémentaires (vers biographie de l'auteur, etc.)
	14b. Possibilité de liens vers d'autres produits (plus ou moins chers, etc.)
	14c. Possibilités de comparatifs de produits (tableaux comparatifs explicites)

Légende : présentation des dimensions 11, 12, 13 et 14 et de leurs sous-dimensions 11a, 11b, 11c, etc. parmi les 31 dimensions en tout.

LISTE DE DIMENSIONS UTILISATEURS DE SITES INTERNET MARCHANDS
TRANSFORMEE EN SURFACE, AVANT L'ENTRETIEN DE VALIDATION AVEC LE
RESPONSABLE DU WEBMASTERING

Dimensions d'évaluation liées plus particulièrement aux sites Internet marchands :

Grands ensembles	Intitulé et définition des dimensions
<p>Avant l'arrivée sur le site</p>	<p>Notoriété de la marque du magasin ou de l'entreprise (ex : Fnac) ou du site Internet lui-même (ex : Amazon) :</p> <ul style="list-style-type: none"> - Est une méga-dimension : la « bonne » notoriété dépend aussi de la qualité reconnue des produits proposés (marques), de la qualité reconnue des services (ex : 48h de la Redoute ou le S.A.V. de Darty), de la sécurisation effective reconnue (premier partenaire Fianet) car le client associe aussi la notoriété à tout cela. - Une bonne médiatisation (publicités, sponsors,...) participe à la notoriété du site. <p>Facilité d'accès liée aux techniques sur Internet :</p> <ul style="list-style-type: none"> - Facilité de mémorisation du nom et de l'adresse URL, - Bon référencement sur les moteurs de recherche Internet (par paiement de sa place : « pôle position »).
<p>Accueil</p>	<p>Page d'accueil : clarté de la vitrine et présence d'informations qui rassurent :</p> <ul style="list-style-type: none"> - En un clin d'œil, l'usager comprend ce qu'il peut acheter et faire sur le site (clarté des titres, slogans et explications succinctes). Toutes les rubriques de produits et services sont présentes dès la page d'accueil. Ces éléments permettent facilement au client d'identifier la cible du site (ex : rubriques destinées exclusivement aux restaurateurs ou au contraire aux consommateurs peu avertis sur les sites de vins, affichage des marques de produits, des fourchettes de prix...), - D'un point de vue visuel, le client reconnait facilement le site (logo présent sur toutes les pages, pas de changement intempestif du site de la page d'accueil), - Des indices de sécurité sont présents (icône du cadenas, logo de FIANET, adresse de l'entreprise, date de mise à jour,...), - Le positionnement affiché en page d'accueil est confirmé dans les pages ultérieures. <p>Possibilité d'accès et qualité du contenu des conditions générales de vente et de sécurité :</p> <ul style="list-style-type: none"> - Clarté du contenu des informations sur les différents modes de paiement, sur les traitements sécurisés, sur les organismes d'assurance, sur la confidentialité des données. Présence de l'icône du cadenas. - Clarté du contenu des conditions de vente (réductions, conditions de retours de produits, ...), des conditions de livraison (disponibilité, frais et délais d'expédition) et sur les services après vente (maintenance,...) et existence d'informations d'identification de tous les services ou prestataires concernés, - Existence d'informations d'identification pour contacter un vendeur (n° de téléphone,...), - Facilité d'accès à toutes ces informations avant de parvenir au processus de commande (liens sur la page d'accueil et sur toutes les pages de présentation de produits). <p>Existence et qualité d'un contact avec un vendeur (aide à l'achat) :</p> <ul style="list-style-type: none"> - Possibilité de contacter un vendeur, un conseiller, - Efficacité du service-vendeur (facile à contacter, personnel disponible, agréable, donnant des conseils appropriés).

Offre – Position- nement du site	<p>Diversité et qualité des produits proposés et de leurs prix :</p> <ul style="list-style-type: none"> - Large choix de familles de produits (Livres, CD, Hifi,...) et diversité dans ces familles ou dans une gamme de produits (ex : toutes les marques et modèles d'appareils photos), - Originalité des produits ou services proposés (raretés, produits vendus seulement à l'étranger, location de vidéos avec pizzas,...), - Qualité des produits (marques, origine), - Rapport qualité/prix des produits bon marché. <p>Qualité des promotions, des nouveautés et existence de divertissements :</p> <ul style="list-style-type: none"> - Qualité des propositions de promotions (inventus, dernières offres, prix réduits intéressants, soldes en rapport avec la saison : ex. sur le linge blanc pour le mois du blanc – en lien avec le choix de positionnement du site, ex. Cdiscount : exclusivités de promotions), - Qualité des propositions de nouveautés (les plus vendues, en rapport avec l'actualité : ex. DVD du dernier film sorti) et de produits en pré-commande (en avant- première), - Possibilités de jeux, de gagner des bons d'achat, des cadeaux sans achat préalable. <p>Différences entre un site Internet et les modes de vente habituels (choix de prix, produits, promotions) :</p> <ul style="list-style-type: none"> - Mêmes promotions, nouveautés, gammes de produits que dans les magasins de la même enseigne (ex. site de la Fnac et ses magasins), - Et/ou en fonction des magasins concurrents du même domaine de vente (choix de positionnement à réaliser en fonction de celui de tous les concurrents).
Offre – présenta-tion et communica- tion de ce position- nement	<p>Existence et qualité des informations liées aux produits, promotions et nouveautés (sur tout support : fiche-produit, newsletters,...) et des stratégies de communication sur les produits :</p> <ul style="list-style-type: none"> - Dès la page d'accueil, sur les fiches-produit, sur les newsletters et autres supports : qualité des informations du descriptif des produits (toutes les caractéristiques), sur les prix et sur les délais et frais d'expédition (informations claires, complètes, actualisées, simples, imprimables), - Sur les fiches descriptives de chaque produit : existence d'illustrations (photos, animation 3D, extraits de musique), de commentaires (critiques de professionnels, d'internautes), et de possibilité de communiquer son appréciation, - Existence de rubriques de promotions, de nouveautés et de divertissements dès la page d'accueil et sur les premières pages suivantes de présentation des produits, - Pertinence de l'information (pas d'informations inutiles sur les produits de la page d'accueil et des autres pages), - Qualité des stratégies de communication (affichage des prix du site et de l'économie réalisée par rapport aux prix du commerce classique). <p>Possibilité d'accès et qualité de propositions d'autres produits (Cross Selling) et d'informations complémentaires:</p> <ul style="list-style-type: none"> - Possibilités de liens vers d'autres produits liés au choix du client (produits plus ou moins chers, etc.) ou aux choix des autres internautes ayant acheté le même produit, - Possibilités de liens vers des informations complémentaires (bibliographie d'un auteur, descriptif d'un pays, mesures des tailles de vêtements d'enfant,...), - Possibilité de comparatifs de produits (tableaux comparatifs systématiques et explicites), - Existence de ces liens dès la fiche du produit sélectionné et sur le panier. <p>Existence et qualité de propositions d'informations personnalisées pour le client (newsletter, Opt-in, Opt-out) et des fonctions associées sur l'interface :</p> <ul style="list-style-type: none"> - Proposition d'informations "en push" (Opt-out), - Possibilité de demande volontaire d'informations personnalisées (Opt-in : newsletter, spécification de préférences pour recevoir des offres en lien avec ce profil) et/ou de personnalisation du contenu du site par choix des rubriques (rare), - Possibilité de s'enregistrer (pour une commande ultérieure plus rapide, la personnalisation « Bonjour Mme... », des bons d'achats, etc.).

<p>Recherche</p>	<p>Existence de différentes possibilités de recherche de produits (moteur de recherche interne, rubriques ...). Possibilités de recherche :</p> <ul style="list-style-type: none"> - par cheminement de rubriques en sous-rubriques, - par mots-clefs (moteur de recherche interne), - par formulaire multi-choix (ex. pour les voyages : ville, horaires, prix) dès la page d'accueil, - par recherche avancée (combinaison de champs ex. titre+auteur) ou guidée par conseils, - par n° de référence catalogue (s'il existe un catalogue-papier), - possibilité d'accès aux sous-catégories dès la page d'accueil (ex : pops-up sur les menus). <p>Facilité et rapidité de tout le processus pour trouver des produits (merchandising ou rayonnage) :</p> <ul style="list-style-type: none"> - Pertinence du choix des rubriques et des choix proposés dans l'affinement des rubriques, - Efficacité du moteur de recherche et de ses options de recherche (recherche par orthographe approchée ex. pour trouver un produit étranger ou par orthographe précise pour obtenir le plus directement le produit recherché, possibilité de recherche sur plusieurs mots-clefs de façon indissociable pour une plus grande pertinence des résultats trouvés), - Facilité d'utilisation des menus déroulants et multi-choix ou champs (prise en compte des préférences soit pour des entrées libres ou soit pour des menus présélectionnés), - Existence de guidage et « intelligence » de la recherche par moteur ou formulaire interne (ex : propositions de voyages en Afrique du Sud au départ de Paris même si la demande était au départ de Nantes ; propositions d'autres recherches en cas de manque de produits), - Rapidité d'aboutissement de la recherche (3 clics environ). <p>Qualité de présentation des résultats de la recherche :</p> <ul style="list-style-type: none"> - Présentation illustrée, aérée et facilement « naviguable » des listes de résultats (petite photo, repérage et navigation par la numérotation des pages, etc.) - Existence de présélections (présentation illustrée de 3 produits seulement et renvois vers les listes d'autres sélections) et classifications des produits trouvés par famille (si nécessaire).
<p>Commande</p>	<p>Panier : facilité d'ajout des produits au panier et de son utilisation :</p> <ul style="list-style-type: none"> - Existence d'un système de panier et de mise en attente de produits (« liste de souhaits », création de compte sans même avoir acheté) et facilité d'accès (bouton du panier existant sur toutes les pages), - Facilité de la procédure de sélection de produits (ex. saisie de la taille, de la quantité, ...), du suivi du cumul des achats (principe d'apparition du panier : à chaque ajout ou à la fin de la sélection complète des produits) et du retour à la sélection de produits ou de poursuite en commande, - Existence de la possibilité d'achat direct (directement au bon de commande, en 1 clic – adresse et n° de CB pré-enregistrés, par référence catalogue à partir de la page d'accueil). <p>Facilité et rapidité des bons de commande (enregistrement/commande/paiement) :</p> <ul style="list-style-type: none"> - Existence d'explications succinctes et claires sur le déroulement de la procédure, - Existence d'un récapitulatif complet des produits, existence d'un champ d'entrée du code promotion pour une commande par bon de réduction, existence de boutons de validation, - Facilité dans les moyens de saisie (menus et champs) et clarté de la présentation de tous les formulaires, - Existence et facilité du principe d'enregistrement /création de compte (par e-mail ou mot de passe plutôt que par n° de client, meilleure perception du formulaire d'enregistrement tout à la fin de la procédure de commande), - Possibilité et facilité de retours en arrière et d'annulation –valable dès le niveau du panier (possibilité de revenir sur un choix ou nombre de produits ; possibilité logique de sélection de mêmes produits de caractéristiques différentes ; possibilité d'annulation à la fin de la commande), - Rapidité de la procédure (peu de données et nombre d'étapes limité).

	<p>Existence de preuves, de confirmation sur l'opération de commande, de conditions vente/sécurité (tout besoin informationnel lors de la procédure de commande) :</p> <ul style="list-style-type: none"> - Existence d'une facture imprimable et clarté du détail des prix des produits et des frais d'expédition, - Accès ou existence d'informations en simultané sur les conditions de sécurité/vente pendant la procédure de commande (liens à toutes les étapes), - Existence d'un message de validation de l'opération de commande ou d'un premier e-mail rapide de confirmation.
Achat	<p>Existence et qualité de la sécurisation et des modes de paiement :</p> <ul style="list-style-type: none"> - Existence de traitements sécurisés (serveur SSL, code supplémentaire de 3 chiffres à côté du panneau de signature de la CB), d'une assurance par une banque ou un organisme tiers, - Existence de différents modes de paiement (CB, chèque, MasterCard, Visa,... Carte Printemps).
	<p>Existence et qualité de la confidentialité des données personnelles (vie privée) :</p> <ul style="list-style-type: none"> - Non-divulgateur à un tiers des données personnelles (adresse, n° de CB,...), « option cochée par défaut » lors de la commande (pas de démarches supplémentaires à réaliser), - Non-obligation de donner des informations trop privées (ex : nombre d'enfants).
Livraison et service après la vente	<p>Possibilité de suivi de la commande :</p> <ul style="list-style-type: none"> - Possibilité de suivi de sa commande sur un panneau de suivi sur le site ou grâce à la réception d'e-mails aux moments utiles (suivi par l'équipe de gestion des commandes), - Possibilité de pré-commander un produit indisponible et d'être prévenu lors de sa remise en vente (suivi par l'équipe de gestion des commandes), - Facilité d'accès dès la page d'accueil à la possibilité de « suivi de commande ».
	<p>Efficacité dans l'expédition (mise en disponibilité, livraison) :</p> <ul style="list-style-type: none"> - Rapidité de l'envoi de produit (mise en disponibilité et livraison rapides) et efficacité dans les garanties de livraison (respect des délais, à domicile, qualité emballage anti-chocs), - Possibilité du choix de la date/heure, du mode de livraison (ex : Chronoposte) et de l'option-cadeau (avec message personnalisé), - Bon niveau (rapport qualité/prix) des frais de livraison.
	<p>Existence et efficacité des services d'après vente (du livreur, des retours de produits, etc.) :</p> <ul style="list-style-type: none"> - Possibilité de retours de produits et de bénéficier d'un service de maintenance, - Possibilité de contacter un service du site ou les prestataires associés par rapport à la livraison à domicile, aux retours de produits (constat des dégâts) ou à la maintenance, - Efficacité de tous ces services A.V. (facilité pour les contacter, existence d'une personne précise qui suit la commande ou efficacité dans l'organisation de suivi, personnel agréable, facilité des modalités de retours, etc.).

Dimensions d'évaluation génériques, relatives à tout site Internet :

Grands ensembles	Intitulé et définition des dimensions
Incontournables (conditions qui doivent être remplies par tout site Internet)	Aide et messages d'erreur : <ul style="list-style-type: none"> - Efficacité de l'aide en ligne : clarté du contenu et facilité d'accès (ex : avatar accompagnateur), - Quantité mesurée, clarté du contenu des messages d'erreur et emplacement visible (ex : en cas de mauvais remplissage d'un formulaire d'inscription).
	Clarté du vocabulaire : <ul style="list-style-type: none"> - Termes et formulations commun(e)s (pas d'abréviations, de jargon technique) et significatif(ve)s (titres accrocheurs, libellés de commandes parlants, descriptions dans un langage simple), - Symboles facilement compréhensibles (ex : €plutôt que E).
	Rapidité de chargement des pages et facilité de repérage des actions possibles (mise en évidence des liens, des boutons de commande, des arborescences) : <ul style="list-style-type: none"> - Rapidité d'affichage des pages et pages proposées effectivement en ligne, - Mise en évidence graphique des commandes et accès principaux (liens, boutons du panier, de validation de commande, des champs de moteur de recherche,...), - Emplacements habituels et constance des accès à travers les pages (liens et barres de navigation habituel(le)s et constant(e)s), - Existence de repérage des arborescences et de possibilité de retours en arrière (à chaque étape).
	Qualité de la présentation (structure, mise en évidence, lisibilité) des informations écrites descriptives en général (page d'accueil et autres pages) : <ul style="list-style-type: none"> - Quantité mesurée d'informations sur chaque page (aération), - Mise en évidence de l'information graphiquement (structure = existence de cadres d'informations ; titre en gras, rubriques de différentes couleurs, informations importantes en couleur, etc.), - Lisibilité visuelle de l'information (taille de police, couleur des caractères / couleur de fond), - Organisation syntaxique de l'information (information importante en début de page ou de paragraphe, bonne catégorisation de l'information, une idée par paragraphe), - Organisation des niveaux d'information (pas trop d'ascenseurs, ni de multi-fenêtrage).
	Esthétique du site : <ul style="list-style-type: none"> - Agrément de l'apparence graphique (agrément et attractivité, ex : apparence ludique ou plutôt sérieuse), - Agrément et non-gêne des animations et autres fioritures (ex : gif animée, personnages et scènes animés), - Existence et agrément d'une musique de fond ou de bruitages (une voix agréable qui annonce les produits ou services).
	Modération et qualité des publicités : <ul style="list-style-type: none"> - Quantité mesurée d'annonces publicitaires et commerciales (annonces visuelles sur le site), - Qualité des annonces publicitaires et commerciales (dans l'apparence graphique, l'emplacement ; thèmes en rapport avec ceux du site).
	Mise à jour du site : <ul style="list-style-type: none"> - Existence de mise à jour fréquente de la présentation graphique (en rapport avec la saison, ex : décoration de Noël), - Existence de mise à jour fréquente du contenu (prix, produits, promotions).
	Compatibilité avec les navigateurs courants : <ul style="list-style-type: none"> - Utilisation possible des navigateurs les plus courants (avec des logiciels intégrés adéquats notamment pour les visuels), - Existence de solutions de repli.

LISTE DE DIMENSIONS UTILISATEURS DE SITES INTERNET MARCHANDS TRANSFORMEE EN SURFACE, APRES L'ENTRETIEN DE VALIDATION AVEC LE RESPONSABLE DU WEBMASTERING

Seulement, des transformations très légères (au niveau du choix des mots ou des exemples dans les définitions) ont été réalisées.

Dimensions d'évaluation liées plus particulièrement aux sites Internet marchands :

Grands ensembles	Intitulé et définition des dimensions
Avant l'arrivée sur le site	<p>Notoriété de la marque du magasin ou de l'entreprise (ex : Fnac) ou du site Internet lui-même (ex : Amazon) :</p> <ul style="list-style-type: none"> - Est une « méga-dimension » : dépend aussi de la qualité reconnue des produits proposés, de la qualité reconnue de la livraison et autres services, de la sécurisation effective. - Bonne médiatisation (publicités, sponsors, ...).
	<p>Facilité d'accès liée aux techniques du web :</p> <ul style="list-style-type: none"> - Facilité de mémorisation du nom et de l'adresse URL, - Bon référencement sur les outils de recherche du web (par positionnement payant pour être en « pôle-position » ou optimisation sur les mots-clefs et contenus).
Accueil	<p>Page d'accueil : clarté de la vitrine et présence d'informations qui rassurent :</p> <ul style="list-style-type: none"> - En un clin d'œil, l'usager comprend ce qu'il peut acheter et faire sur le site. Les éléments de la page d'accueil permettent facilement au client d'identifier la cible du site, - D'un point de vue visuel, le client reconnait facilement le site, - Des indices de sécurité sont présents (FIANET, CNIL), - Le positionnement affiché en page d'accueil est confirmé dans les pages ultérieures.
	<p>Possibilité d'accès et qualité du contenu des conditions générales de vente et de sécurité :</p> <ul style="list-style-type: none"> - Clarté du contenu des informations sur les différents modes de paiement, la sécurisation, les organismes d'assurance, la confidentialité, les conditions de vente (modalités retours produits), de livraison, sur les services après vente (maintenance ou autre) ou le service-vendeur et existence d'informations d'identification de tous les services ou prestataires concernés (n° tél, e-mail), - Facilité d'accès à toutes ces informations avant de parvenir au processus de commande.
	<p>Existence et qualité d'un contact avec un vendeur (aide à l'achat) :</p> <ul style="list-style-type: none"> - Possibilité de contacter un vendeur, un conseiller, - Efficacité du service-vendeur (facile à contacter, donnant des conseils appropriés).
Offre – Positionnement du site	<p>Diversité et qualité des produits proposés et de leurs prix :</p> <ul style="list-style-type: none"> - Large choix de familles de produits et diversité dans une gamme de produits, - Originalité des produits ou services proposés (raretés) et/ou qualité des produits (marques), - Rapport qualité/prix des produits bon marché.
	<p>Qualité des promotions, des nouveautés et existence de divertissements :</p> <ul style="list-style-type: none"> - Qualité des propositions de promotions (soldes en rapport avec la saison : ex. sur le linge blanc pour le mois du blanc), - Qualité des propositions de nouveautés et de produits en pré-commande (en avant 1^{ière}), - Possibilités de jeux, de gagner des bons d'achat, des cadeaux.

	<p>Différences entre un site Internet et les modes de vente habituels (choix de prix, produits, promotions) :</p> <ul style="list-style-type: none"> - Mêmes promotions, nouveautés, gammes de produits, etc. que dans les magasins de la même enseigne (ex : pour la Fnac) ou en fonction des magasins concurrents du même domaine de vente (ex : pour un site comme Alapage qui vend des livres sans avoir de magasins) car les clients font des comparaisons.
Présentation /communication /valorisation de ce positionnement	<p>Existence et qualité des informations liées aux produits, promotions et nouveautés (sur tout support : fiche-produit, newsletters,...) et des stratégies de communication sur les produits :</p> <ul style="list-style-type: none"> - Dès la page d'accueil, sur les fiches-produit, sur les newsletters et autres supports : qualité des informations du descriptif des produits (caractéristiques), sur les prix et les délais et frais d'expédition, existence d'illustrations, de commentaires divers et de possibilité de communiquer son appréciation, - Existence de rubriques de promotions, de nouveautés dès la page d'accueil, - Qualité des stratégies de communication (affichage de l'économie réalisée / prix classique), - Pertinence de l'information (pas inutile).
	<p>Possibilité d'accès et qualité de propositions d'autres produits (Cross Selling) et d'informations complémentaires:</p> <ul style="list-style-type: none"> - Possibilités de liens vers d'autres produits liés au choix du client ou aux choix des autres internautes, vers des informations complémentaires (ex. discussion en ligne), de comparatifs de produits, - Existence de ces liens dès la fiche du produit sélectionné et sur le panier.
	<p>Existence et qualité de propositions d'informations personnalisées pour le client (newsletter, Opt-in, Opt-out) et des fonctions associées sur l'interface :</p> <ul style="list-style-type: none"> - Proposition d'informations "en push" (Opt-out), - Possibilité de demande volontaire d'informations personnalisées (Opt-in, ex. newsletter avec personnalisation des rubriques), - Possibilité de s'enregistrer (pour une commande ultérieure plus rapide, la personnalisation).
Recherche	<p>Existence de différentes possibilités de recherche de produits (moteur de recherche interne, rubriques ...) :</p> <ul style="list-style-type: none"> - Possibilités de recherche par rubriques et sous-rubriques (le rayonnage des produits), par mots-clefs, par formulaire multi-choix dès la page d'accueil, par recherche avancée ou guidée par conseils, par n° de référence catalogue (ex. catalogue La Redoute), - Possibilité d'accès aux sous-catégories dès la page d'accueil (ex : menus déroulants en pops-up).
	<p>Facilité et rapidité de tout le processus de recherche de produits (merchandising ou rayonnage) :</p> <ul style="list-style-type: none"> - Pertinence du choix des rubriques et sous-rubriques, - Efficacité du moteur de recherche et de ses options de recherche (recherche par orthographe approchée ou précise, exclusions de mots-clefs non pertinents), - Facilité d'utilisation des menus déroulants et multi-choix ou champs, - Existence de guidage et « intelligence » de la recherche par moteur ou formulaire interne dépendamment des critères demandés (ex.. proposition des voyages en Afrique du Sud au départ de Paris même si la demande était au départ de Nantes, pas de message « désolés, nous ne pouvons pas ... »), - Rapidité d'aboutissement de la recherche (3 clics environ).
	<p>Qualité de présentation des résultats de la recherche :</p> <ul style="list-style-type: none"> - Présentation illustrée, aérée et facilement « naviguable » des listes de résultats, - Existence de présélections (lien de renvois sur des listes plus exhaustives) et classifications des produits trouvés par famille (ex. catégorisation par documentaires, romans, etc. pour la demande de thèmes généraux comme « La Loire » dans les livres).
Commande	<p>Panier : facilité d'ajout des produits au panier et de son utilisation :</p> <ul style="list-style-type: none"> - Existence d'un système de panier et de mise en attente de produits (« liste de souhaits ») et facilité d'accès, - Facilité de la procédure de sélection de produits (ex. taille, quantité,...), du suivi du cumul des achats et du retour à la sélection de produits ou de poursuite en commande, - Possibilité d'achat direct (en 1 clic, par référence catalogue sur le bon de commande).

	<p>Facilité et rapidité des bons de commande (enregistrement/commande/paiement) :</p> <ul style="list-style-type: none"> - Existence d'explications succinctes, - Existence d'un récapitulatif complet des produits, de boutons de validation, de tous les champs nécessaires liés au choix du mode de livraison (éventuellement date/heure) ou du mode de paiement, - Facilité dans les moyens de saisie et clarté de la présentation de tous les formulaires, - Existence et facilité du principe d'enregistrement /création de compte, - Possibilité et facilité de retours en arrière et d'annulation (d'un produit ou d'une opération de commande), - Rapidité de la procédure (peu de données et nombre d'étapes limité).
	<p>Existence de preuves, de confirmation sur l'opération de commande et des conditions vente/sécurité (tout besoin informationnel lors de le procédure de commande) :</p> <ul style="list-style-type: none"> - Existence d'une facture imprimable et clarté du détail de tous les frais, d'informations en simultané sur les conditions de sécurité/vente, d'un message de validation de l'opération de commande ou d'un premier e-mail rapide de confirmation.
Achat	<p>Existence et qualité de la sécurisation et de différents modes de paiement :</p> <ul style="list-style-type: none"> - Existence de traitements sécurisés (ex. SSL ou code à 3 chiffres derrière CB), d'une assurance par un organisme tiers, - Existence de différents modes de paiement (CB, chèque).
	<p>Existence et qualité de la confidentialité des données personnelles (vie privée) :</p> <ul style="list-style-type: none"> - Non-divulgation à un tiers des données personnelles, - Non-obligation de donner des informations trop privées (ex : nombre d'enfants).
Livraison et service après la vente	<p>Possibilité suivi de la commande :</p> <ul style="list-style-type: none"> - Possibilité de suivi de sa commande (panneau de suivi, e-mails) et qualité des modalités de suivi de commande en ligne (modification de l'adresse, annulation, etc.) - Possibilité de réserver un produit indisponible (système d'alertes pour le client de remise en vente), - Facilité d'accès dès la page d'accueil à la possibilité de « suivi de commande ».
	<p>Efficacité dans l'expédition (mise en disponibilité, livraison) :</p> <ul style="list-style-type: none"> - Rapidité et efficacité dans les garanties de livraison (ex. qualité emballage anti-chocs), - Possibilité du choix de la date/heure (surtout pour les produits consommables), du mode de livraison (ex. Chronoposte) et de l'option cadeau, - Bon niveau (rapport qualité/prix) des frais de livraison.
	<p>Existence et efficacité des services d'après vente (problème de paiement, du livreur, des retours de produits, etc.) :</p> <ul style="list-style-type: none"> - Possibilité de retours de produits et de bénéficier d'un service de maintenance, - Possibilité de contacter un service ou les prestataires associés par rapport à la livraison à domicile, à un problème de paiement et autres problèmes , - Efficacité de tous ces services (facilité pour les contacter et des modalités de retours, efficacité de la réponse..).

Dimensions d'évaluation génériques, relatives à tout site Internet :

Grands ensembles	Intitulé et définition des dimensions
Incontournables (conditions qui doivent être remplies par tout site Internet)	Aide et messages d'erreur : <ul style="list-style-type: none"> - Efficacité de l'aide en ligne - clarté du contenu et facilité d'accès (ex : avatar accompagnateur dans la mesure des possibilités techniques liées à l'équipement des usagers-cibles), - Quantité mesurée, clarté du contenu des messages d'erreur et emplacement visible (ex : en cas de mauvais remplissage d'un formulaire d'inscription).
	Clarté du vocabulaire : <ul style="list-style-type: none"> - Termes et formulations parlant(e)s, pas d'abréviation, vocabulaire simple, - Symboles facilement compréhensibles (ex : € plutôt que E, dans la mesure des possibilités techniques liées à l'équipement des usagers-cibles).
	Rapidité de chargement des pages et facilité de repérage des actions possibles (mise en évidence des liens, des boutons de commande, des arborescences) : <ul style="list-style-type: none"> - Rapidité d'affichage et pages proposées effectivement en ligne, - Mise en évidence graphique des commandes et accès principaux, - Emplacements habituels et constance des accès à travers les pages (barres de navigation), - Existence de repérage des arborescences et de possibilité de retours en arrière.
	Qualité de la présentation des informations écrites ou descriptives (page d'accueil et autres pages) : <ul style="list-style-type: none"> - Quantité mesurée d'informations (aération), - Mise en évidence de l'information graphiquement (structurée, rubriques de différentes couleurs, fond blanc, etc.) et lisibilité (taille de police), - Organisation syntaxique de l'information (information importante en début de page), - Organisation des niveaux d'information (pas de multiples ascenseurs, ni trop de multi-fenêtrage).
	Pertinence des informations écrites ou descriptives (page d'accueil et autres pages) : <ul style="list-style-type: none"> - Pertinence et utilité de l'information (information concise mais suffisante, en lien avec la tâche que l'utilisateur doit effectuer).
	Esthétique du site : <ul style="list-style-type: none"> - Agrément de l'apparence graphique (joli, attractivité, style général), - Agrément et non-gêne des animations et autres fioritures, - Existence et agrément d'une musique de fond et/ou d'une voix accompagnatrice, dans la mesure des possibilités techniques liées à l'équipement des usagers-cibles.
	Modération et qualité des publicités : <ul style="list-style-type: none"> - Quantité mesurée d'annonces publicitaires et commerciales et qualité des annonces publicitaires et commerciales (apparence graphique, emplacement, thèmes en rapport avec ceux du site).
	Mise à jour du site : <ul style="list-style-type: none"> - Existence de mise à jour fréquente de la présentation graphique (ex : décoration de Noël) et du contenu dans la base de données (produits, promotions) ou dans le contenu éditorial.
Compatibilité avec les équipements courants : <ul style="list-style-type: none"> - Utilisation possible de tous les navigateurs et des logiciels les plus courants (notamment pour les visuels) et de toutes les plates-formes (P.C., Mac Intosh) et existence de solutions de repli (ex. logiciels qui transforment l'image en texte). 	

ANNEXES DU CHAPITRE 5

LES REGROUPEMENTS DES DIMENSIONS UTILISATEURS DES SITES INTERNET MARCHANDS PAR LE PREMIER PROFESSIONNEL DE LA VEILLE-MARKETING (SPECIALISTE DES SITES MARCHANDS)

(pages suivantes)

Figure 18 : Regroupements des dimensions utilisateurs par le professionnel de la veille-marketing, spécialiste des sites marchands, lors de l'entretien avec utilisation du Sitemarchand.com

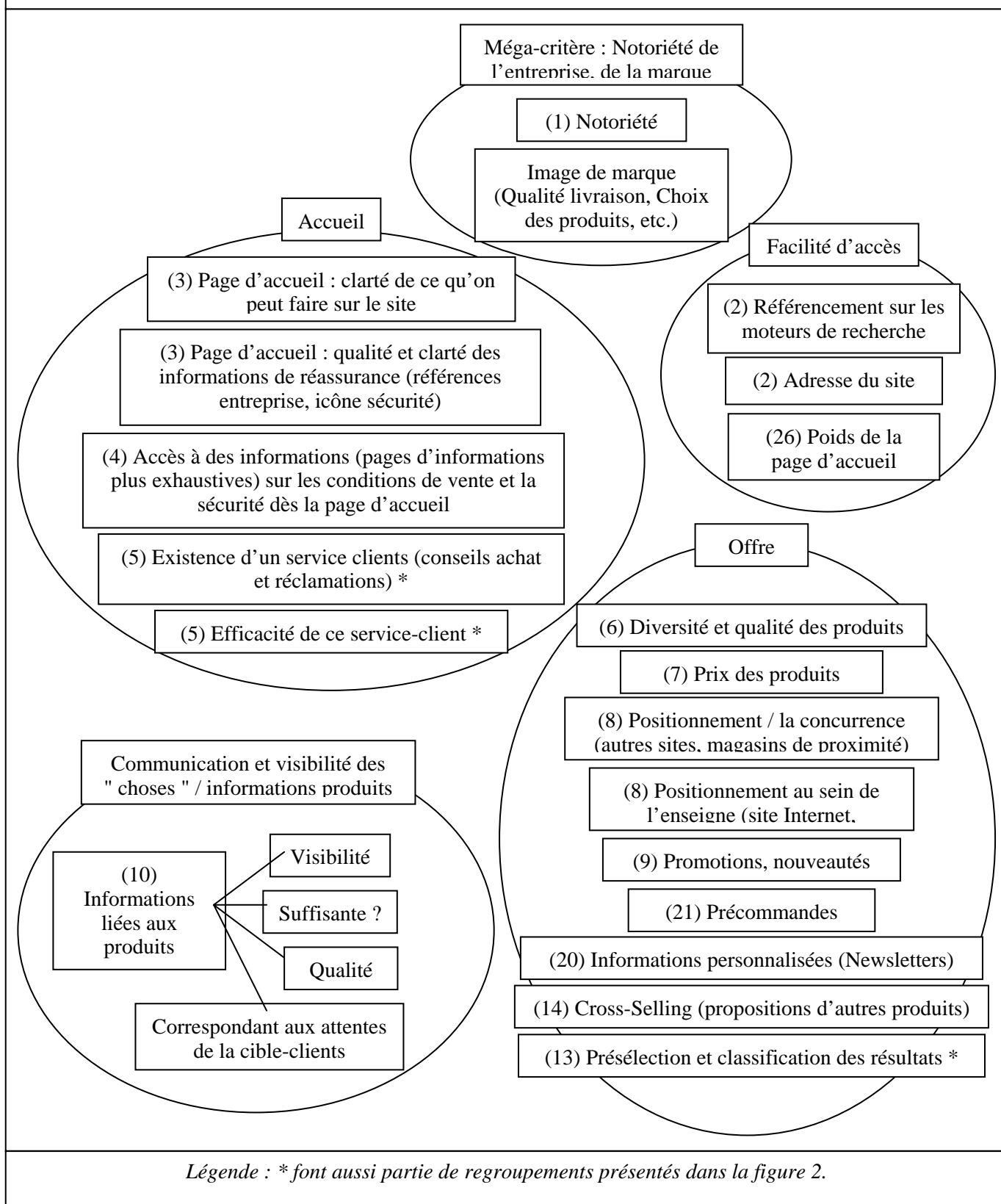
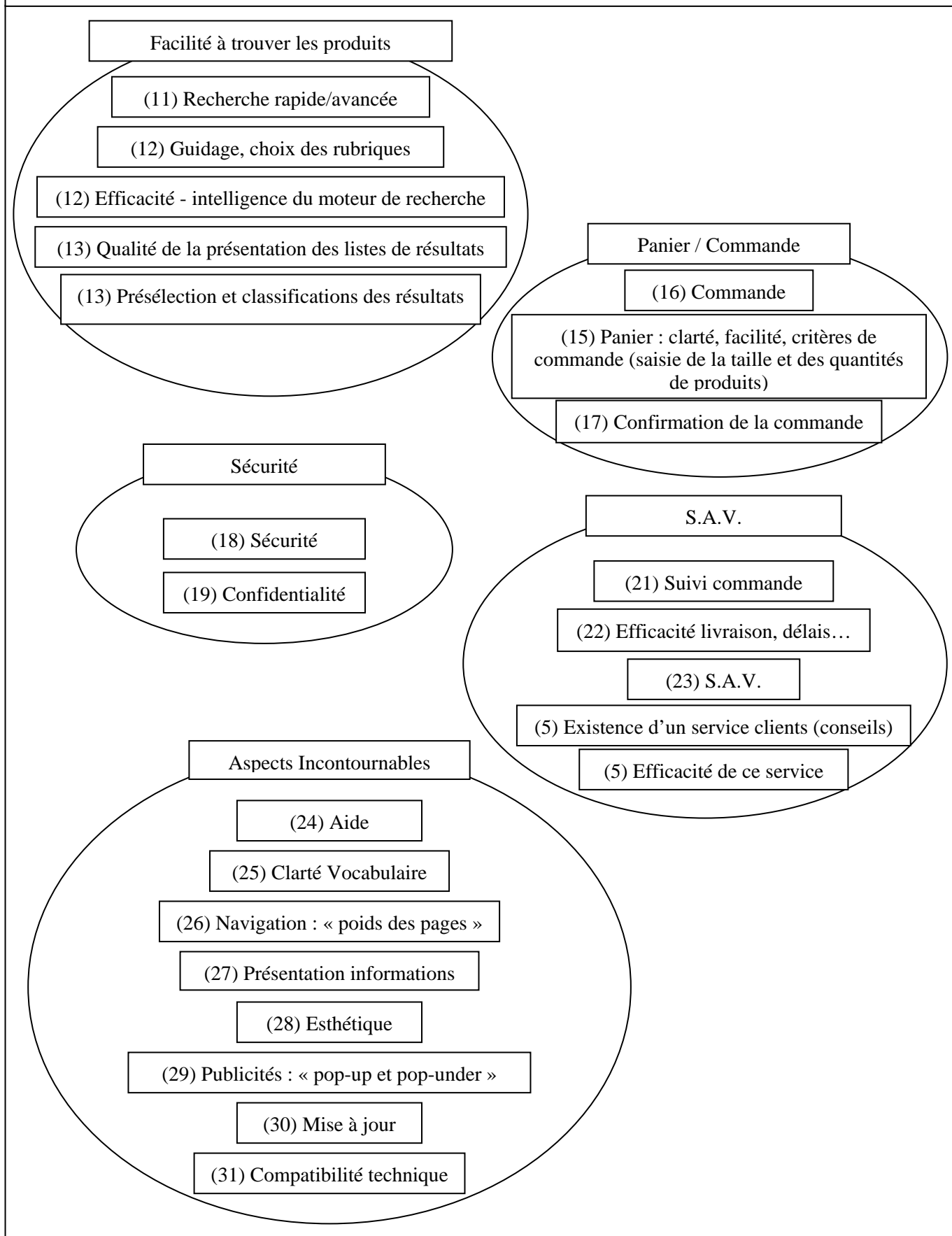


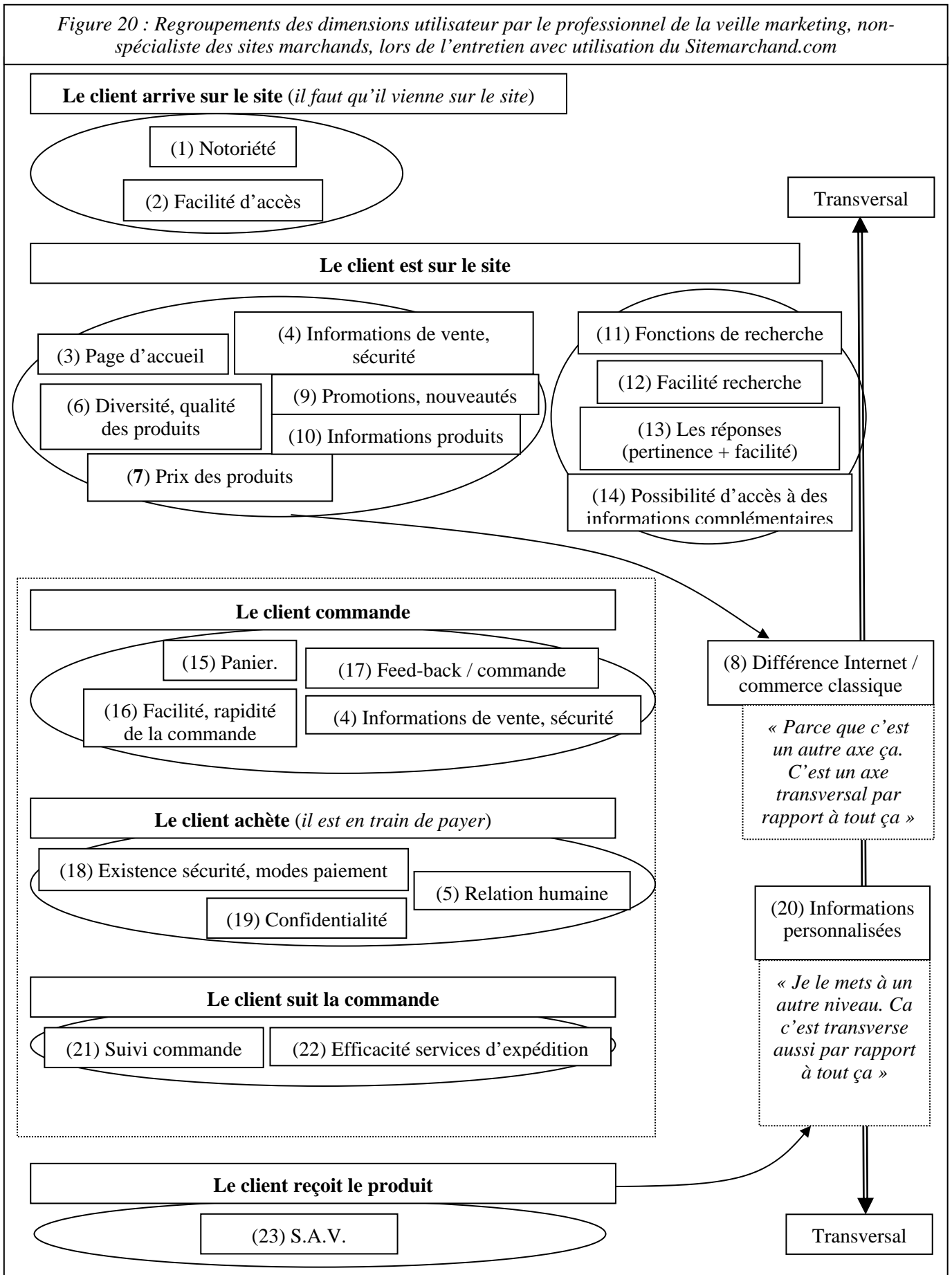
Figure 19 : Regroupements des dimensions utilisateur par le professionnel de la veille marketing, spécialiste des sites marchands, lors de l'entretien avec utilisation du Sitemarchand.com



LES REGROUPEMENTS DES DIMENSIONS UTILISATEURS DES SITES INTERNET MARCHANDS PAR LE SECOND PROFESSIONNEL DE LA VEILLE-MARKETING (NON SPECIALISTE DES SITES MARCHANDS)

(page suivante)

Figure 20 : Regroupements des dimensions utilisateur par le professionnel de la veille marketing, non-spécialiste des sites marchands, lors de l'entretien avec utilisation du Sitemarchand.com



ANNEXES DU CHAPITRE 6

RECAPITULATIF DES « BESOINS EN EVALUATION » AVANT ET APRES VALIDATION DE L'EBAUCHE DE L'OUTIL D'AIDE A LA CONCEPTION D'EVALUATIONS

Voici un récapitulatif de tous les « besoins en évaluation » avant et après validation de l'ébauche de l'outil d'aide à la conception d'évaluations, présenté dans le nouvel ordre (celui obtenu après validation). La colonne de tableaux de gauche correspond aux besoins en évaluation avant les entretiens de validation avec les professionnels du marketing et de la conception, tandis que la colonne de droite correspond aux besoins en évaluation après ces entretiens de validation : certains sont supprimés, d'autres déplacés, d'autres encore ajoutés et enfin quelques uns voient leur intitulé reformulé. Il s'agit donc là d'un comparatif illustrant l'évolution de l'ébauche à la boîte à outils finale. Les éléments en grisé dans les tableaux de la colonne de droite correspondent aux nouveaux besoins en évaluation (en gris foncé) et aux besoins en évaluation modifiés (en gris clair ; modification de leur ordre dans la boîte à outils, de leur intitulé ou d'une partie de leur contenu).

« Notoriété perçue », avant et après validation

Besoins en évaluation avant validation
Notoriété
Image de marque
Efficacité de la médiatisation par rapport à la découverte du site
Médiatisation, image de marque et raisons de retour des visiteurs sur le site
Efficacité de la médiatisation « offline » du site sur ses ventes et sa notoriété
Réception des messages publicitaires

Besoins en évaluation après validation
Notoriété
Image de marque
Efficacité et poids relatif des différents supports médiatiques
Raisons de retour des acheteurs sur le site
Réception des messages publicitaires

« Facilité d'accès & Capacité système », avant et après validation

Besoins en évaluation avant validation
Référencement
Simplicité Adresse URL
Rapidité de chargement

Besoins en évaluation après validation
Référencement et liens d'accès au site
Rapidité de chargement
Simplicité Adresse URL

« Page d'accueil », avant et après validation

Besoins en évaluation avant validation
Compréhension immédiate de l'objectif du site
Réassurance
Accès à tous les services et rubriques

Besoins en évaluation après validation
Compréhension immédiate de l'objectif du site
Réassurance
Rubriques & Services principaux

« Look, présentation, visibilité et navigation », avant et après validation

Besoins en évaluation avant validation
Esthétique
Attractivité (ambiance)
Reconnaissance visuelle
Graphismes
Aération
Organisation, Mise en page, Mise en valeur graphique
Visibilité des informations importantes
Visibilité & Emplacements des boutons d'action

Besoins en évaluation après validation
Reconnaissance visuelle
Attractivité & Esthétique
Graphismes & Animations
Aération
Organisation & Mise en valeur graphique
Boutons d'action (visibilité / emplacements)
Apparence des courriels
Apparence des « Cross-Sellings » et des Opérations spéciales en ligne
Publicité en ligne

« Offre & Positionnement », avant et après validation

Besoins en évaluation avant validation
Choix de produits/services
Prix
Promotions & Nouveautés
Positionnement interne (même enseigne)
Positionnement par rapport à la concurrence

Besoins en évaluation après validation
Positionnement par rapport à la concurrence
Les motifs d'achats
Choix de produits/services
Prix
Promotions & Nouveautés
Avantages clients
Positionnement interne

« Description des produits », avant et après validation

Besoins en évaluation avant validation
Caractéristiques produits
Commentaires & Critiques
Illustrations
Annonce prix
Annonce disponibilité & frais d'envois
Mise à jour
Visibilité des informations-produits
Editorial (valeur ajoutée)

Besoins en évaluation après validation
Utilité des informations produits
Caractéristiques descriptives des produits
Commentaires & Critiques
Illustrations
Annonce prix
Annonce disponibilité & frais d'envois
Aspects visuels de toutes les informations produits
Mise à jour – « Fraîcheur » des produits
Editorial (valeur ajoutée)

« Netmarketing », avant et après validation

Besoins en évaluation avant validation
« Cross-Sellings »
Présélections dans les listes de résultats
« Newsletters » (non personnalisées)
« Newsletters » ou offres personnalisées
Offres personnalisées (« Opt-in »)
Courriels en « push » (« Opt-out »)
Publicité en ligne
Avantages clients & Cartes privatives

Besoins en évaluation après validation
Esprit communautaire et offres spéciales ou croisées :
Esprit communautaire
Vote électronique & Témoignages
Présélections dans les listes de résultats
« Cross-Sellings »
Opérations & Boutiques spéciales en ligne
Promotion « online » des produits
Aspects visuels des « Cross-Sellings », des Opérations spéciales et de la publicité
Liens vers des sites d'informations, des chats, des forums
« Newsletters », courriels « en push » et « compte perso » :
Courriels en « push » (Opt-in / Opt-out)
« Newsletters » non personnalisées
« Newsletters » personnalisées
Aspects visuels de tout courriel
Compréhension de la page « compte perso »

« Logistique & Services », avant et après validation

Besoins en évaluation avant validation
Garanties de livraison (Respect délai & Bon état produit)
Emballage cadeau
Frais d'envois & Somme minimum d'achat pour une livraison gratuite
Possibilité de RDV
Relais colis (valeur ajoutée)
Suivi de commande
Retours de produits
Service de maintenance

Besoins en évaluation après validation
Utilité des services
Suivi de commande
Garanties de livraison
Possibilité de RDV & Relais colis
Emballage cadeau
Frais d'envois & Somme minimum d'achat pour une livraison gratuite
Retours de produits
Service de maintenance

« Interactivité », avant et après validation

Besoins en évaluation avant validation
Possibilité de contacter un service clients
Moyens pour contacter le service
Qualité des réponses apportées & Convivialité
Service spécial conseil achat
Liens vers des sites d'informations, des chats, des forums
Vote électronique & Témoignages

Besoins en évaluation après validation
Possibilité de contacter un service clients
Moyens pour contacter le service
Qualité des réponses apportées & Convivialité

« Sécurité & Paiement », avant et après validation

Besoins en évaluation avant validation
Sentiment de sécurité
Non-archivage du n° CB
Partenaires financiers & Organismes de contrôle
Confidentialité des informations personnelles
Facilités de paiement

Besoins en évaluation après validation
Facilités de paiement
Sentiment de sécurité
Confidentialité
Options de sécurisation
Partenaires financiers & Organismes de contrôle

« Facilité à trouver un produit », avant et après validation

Besoins en évaluation avant validation
Trouver le produit
Recherche par mots-clefs
Les rubriques et sous-rubriques
Listes de résultats

Besoins en évaluation après validation
Trouver le produit
Recherche par mots-clefs
Rubriques et sous-rubriques de produits
Recherche avancée
Autres modes de recherche
Listes de résultats
Boutons d'action – visibilité - navigation

« Facilité à commander un produit », avant et après validation

Besoins en évaluation avant validation
Commander le produit
Saisie de la taille, de la quantité des produits
Panier
Bon de commande
Récapitulatif de produits & frais d'envoi

Besoins en évaluation après validation
Commander le produit
Saisie des caractéristiques des produits
Panier
Bon de commande
Récapitulatif frais & délais
Création d'un identifiant ou « compte perso »
Boutons d'action – visibilité - navigation

« Informations associées au service de vente », avant et après validation

Besoins en évaluation avant validation
Qualité & Choix de ces informations
Complétude de ces informations
Compréhension
Visibilité de toutes ces informations

Besoins en évaluation après validation
Pertinence & Compréhension de ces informations
Complétude de ces informations
Visibilité de toutes ces informations

Cette thèse se situe dans le champ de l'évaluation ergonomique de produits et de services. Une bonne évaluation doit répondre à deux types de critères qui ne sont pas nécessairement convergents : d'une part, ceux relatifs à la production de résultats scientifiquement valides et, d'autre part, ceux relatifs à l'utilisation des résultats par les professionnels destinataires. Nous explorons les deux dimensions. Nous produisons des connaissances sur la nature de la production verbale et la productivité de deux techniques utilisées en vue de la construction d'un questionnaire d'évaluation dans le domaine du commerce électronique : l'entretien individuel et le focus group. Par ailleurs, nous construisons un outil d'aide à l'élaboration d'évaluations pour des professionnels du commerce électronique qui produisent des résultats et les utilisent pour transformer leur site Web marchand.

Ergonomic Evaluation of Products and Services: the E-Commerce Example

This Ph.D. deals with the ergonomic evaluation of products and services. A good evaluation must meet two types of criteria, which do not necessarily converge: on the one hand, those concerning the production of scientifically sound results and, on the other hand, those concerning the use of the results by addressees. Both issues are explored. Knowledge is produced concerning the nature of the verbal content produced and the performance of the two techniques used in order to construct an evaluation questionnaire in the electronic commerce field: the in-depth one-on-one interview and the focus group. Furthermore, a tool is constructed which aims to help electronic commerce professionals, who produce results and use them to design their commercial website, to draw up evaluations.

ERGONOMIE

MOTS-CLES : évaluation ergonomique, entretien individuel, focus group, technique d'élicitation, enquête par questionnaire, outil d'aide à l'évaluation, commerce électronique, site Internet marchand.

Equipe "C3U", Laboratoire "Paragraphe" EA 349, UFR de Psychologie
Université Paris 8, 2, rue de la liberté, 93526 Saint Denis Cedex.