Réflexions méthodologiques autour du Napping : vers une intégration du comportement du sujet dans l'analyse des données de Napping

Minh Tam Le

N⁰ ordre :
N⁰ série :

THESE \ AGROCAMPUS OUEST

Sous le sceau de l'Université Européenne de Bretagne

Pour obtenir le diplôme de :

DOCTEUR DE L'INSTITUT SUPERIEUR DES SCIENCES AGRONOMIQUES, AGRO-ALIMENTAIRES, HORTICOLES ET DU PAYSAGES

Spécialité : Mathématiques Appliquées

École Doctorale : Vie-Agro-Santé

Présentée par:

Tâm Minh LÊ

Réflexions méthodologiques autour du Napping® : vers une intégration du comportement du sujet dans l'analyse des données de Napping®

Soutenue le 21 Septembre 2015 devant la commission d'Examen :

Composition du jury:

Julien DELARUE	AgroParisTech, France	Rapporteur
Sylvie CHOLLET	Institute Supérieur d'Agriculture de Lille, France	Rapporteur
Dzung Hoang NGUYEN	HCMC University of Technology, Vietnam	Examinateur
Thomas Croguennec	Agrocampus Ouest, France	Examinateur
Sébastien LÊ	Agrocampus Ouest, France	Directeur
François HUSSON	Agrocampus Ouest, France	Directeur

(This page is left blank)

Remerciements

La thèse est une parenthèse importante pour les jeunes chercheurs qui veulent faire de la recherche de façon autonome, comme un enfant veut marcher seul dans la rue. Pour pouvoir marcher dans ma future rue, je tiens à remercier tous ceux qui m'ont guidé et ceux qui m'ont accompagné depuis mes premières pas.

Je tiens à remercier en premier lieu les membres de mon jury de thèse. Premièrement, Monsieur Julien Delarue, maître de conférences à AgroParisTech, et Madame Sylvie Chollet, maître de conférences à l'ISA Lille, pour leur participation en qualité de rapporteur de mon travail, pour le temps consacré à la lecture de ce rapport, et pour les remarques judicieuses qu'ils m'ont indiquées. Deuxièmement, Monsieur Thomas Croguennec, Professeur à Agrocampus Ouest, pour l'intérêt qu'il a manifesté en qualité de président du jury. Enfin, Monsieur Dzung Hoang Nguyen, Professeur à l'IPHCM pour sa participation en qualité d'examinateur d'une part, et pour ses enseignements stimulants dans le domaine de l'analyse sensorielle depuis mes débuts d'autre part.

Je tiens à remercier en deuxième lieu mes deux directeurs de thèse. Tout d'abord, François Husson, pour m'avoir fait confiance malgré les connaissances plutôt légères que j'avais en statistique, puis pour m'avoir conseillé. Ensuite, j'adresse de chaleureux remerciements à Sébastien Lê, non seulement un directeur de thèse mais aussi un frère de cœur qui m'a accompagné et encouragé tout au long de ma formation et ma vie en France.

Je remercie chaleureusement toute l'équipe du Laboratoire de Mathématique Appliquées d'Agrocampus Ouest (LMA2). Tout d'abord, Monsieur Jérôme Pagès, ancien directeur, ainsi que Monsieur David Causeur, actuel directeur pour m'avoir accueilli au sein de cette institution, et pour les conseils que j'ai eu l'honneur de recevoir de leur part. Je souhaite ensuite remercier les secrétaires Karine Bagory et Elisabeth Lenauld pour avoir géré le côté administratif de ma thèse. Je n'oublie pas Marie Verbanck qui m'a guidé dans l'utilisation de Latex et Ubuntu dès mes premiers jours au Laboratoire. Un grand merci à mon cher ami Vincent Audigier, avec qui j'ai partagé un bureau pendant ces trois années, qui était toujours disponible en tous lieux pour les corrections « des fautes en français », pour le café et pour un morceau de guitare. Je n'oublie pas la gentillesse d'Emeline Perthame qui m'a aidé à répéter ma soutenance avec patience, et qui a souvent accepté de prendre une pause en marchant dans le couloir. Je tiens à remercier Margot Brard qui a sacrifié tout un weekend pour construire le diaporama avant ma soutenance. Je ne pourrais terminer ces remerciements sans penser à Julie Josse, Magalie Houée-Bigot et Mathieu Emily pour leur bonne humeur et leurs encouragements quotidiens au laboratoire.

Je remercie bien sûr tous les étudiants en master qui, au cours de ces trois années, m'ont permis de travailler sur des données réelles : Quentin Carré, Dzung Quoc Nguyen et Samuel Quettier pour les données sur les émotions, Sandra Jourdan et Solène Morvant pour les données des parfums et des vidéos.

Je voudrais également remercier tous mes collègues Uyen Xuan Thi Phan, Thanh Ba Nguyen et Khuong Thanh Nguyen. Mes conversations avec vous ont toujours été enrichissantes aussi bien sur le plan scientifique que sur la vie courante. J'en profite pour remercier Emmanuel Dollé pour son aide à construire le serveur du e-Napping, et Gabriel Jalam pour m'avoir permis d'utiliser la plateforme e-enquête.

À titre plus personnel, je remercie mes parents qui m'ont toujours encouragé, ma femme Linh Thuy Lê et mon fils Tri Minh Lê qui ont su être patients pendant ces trois années. Je sais que mon absence a été longue et je vous dois beaucoup. Enfin, Tâm ne serais rien sans R.

Table of contents

Résumé

La pratique de l'analyse sensorielle a connu une révolution au cours de ces quinze dernières années. Une frontière a été franchie lorsqu'il s'est agi de faire participer des sujets non entrainés au processus d'évaluation sensorielle, jusqu'alors réservé aux seuls sujets entrainés.

Pour s'adapter à la nature du sujet, de nouvelles méthodes de recueil de données sensorielles sont apparues, parmi lesquelles on peut citer le tri libre et ses différentes déclinaisons (Chollet, Lelièvre, Abdi, & Valentin, 2011) ou le Napping® (Pagès, 2005).

Alors que le tri libre est une méthode pour laquelle la communauté scientifique, notamment les psychologues, a un certain recul, relativement peu d'articles méthodologiques ont été écrits sur le Napping®. Pourtant, cette méthode connaît un engouement fort, comme peut en témoigner le nombre relativement important de publications portant sur des applications du Napping®.

L'objectif de ce travail de thèse est d'amener des éléments de réflexion méthodologiques autour du Napping®. Le premier chapitre de ce manuscrit rappelle au lecteur ce qu'est le Napping®, à travers une description de la tâche et des données qu'elle fournit, telles qu'elles ont été imaginées initialement par Jérôme Pagès. C'est dans ce même chapitre que le lecteur peut se rendre compte de l'importance de prendre en compte le comportement du sujet lors de la tâche, tant pour la compréhension des données que pour leur traitement. Le deuxième chapitre de ce manuscrit propose un modèle du comportement du sujet dans le cas où le nombre de stimuli est relativement élevé, une situation couramment rencontrée en pratique. Ce modèle et ses limites amènent le lecteur au chapitre 3, dans lequel il est présenté une méthode originale de recueil de données de Napping® appelée digit-tracking. Cette méthode permet d'observer le comportement du sujet au cours du temps par l'enregistrement du positionnement relatif des stimuli tout au long de la tâche et non plus seulement lors de l'étape finale. Enfin le dernier chapitre du manuscrit présente l'outil qui a été développé au cours de cette thèse et qui permet de recueillir des données de Napping® au cours du temps. Cet outil est en fait une plateforme collaborative couplée à un logiciel de recueil de données sur tablette tactile, qui permet aux chercheurs de partager une partie ou l'ensemble de leurs résultats.

Abstract

The practice of sensory analysis has undergone a tremendous revolution during the past fifteen years. A boundary was established when it came to involve untrained subjects in sensory evaluation process, which was previously reserved only for trained subjects.

In order to be adaptable to the natural cognitive process of the subject, new methods for collecting sensory data have been developed, including the free sorting task and its many different forms (Chollet, Lelièvre, Abdi, & Valentin, 2011) or Napping® (Pagès, 2005).

While many studies have been focused on the methodology of the free sorting task, few methodological articles have been written on the Napping®. Nevertheless, this method has a great significance as we can cite a large numbers of publications regarding to its applications.

The objective of this dissertation aims to bring the methodological reflection elements around the Napping® method. The first chapter provides the concept of Napping® via a description (i.e., an explanation) of its procedure and data format, as it was originally conceived by Jérôme Pagès. This chapter also discussed the importance of taking into account the subject's behaviour during the task for a better understanding of the data collected and their analysis. The second chapter proposes a model of the subject's behaviour in which the number of stimuli is relatively high, a situation commonly encountered in practice. This model and its limitations lead to Chapter 3, where a new technique for collecting Napping® data, named digit-called tracking, is presented. This new technique of data collection allows observing the subject's behaviour over time by means of recording the relative positioning of all stimuli throughout the task, rather than at its final step. Finally, the last chapter presents a tool devoted to collecting Napping® data over time. This tool is a collaborative platform that works together with an application for collecting data on tactile tablet devices, which enables researchers to share totally or partially their study resources.

Avant-propos

Ce travail de thèse a été conduit au sein de l'unité pédagogique "Mathématiques Appliquées" (Département Statistique et Informatique) d'Agrocampus Ouest (Rennes, France) sous la direction de François Husson et Sébastien Lê.

Cette unité de l'Institut Supérieur des Sciences Agronomiques, Agroalimentaires, Horticoles et du Paysage a une activité de recherche centrée sur le développement de nouvelles méthodologies en statistique appliquée. Les applications privilégiées du laboratoire concernent les domaines de la Génomique et de l'Analyse Sensorielle. C'est dans ce dernier domaine que se positionne ce travail de thèse.

En 2005, Jérôme Pagès, alors directeur de l'unité, propose une nouvelle méthodologie de recueil de données sensorielles appelée Napping®. Un des objectifs principaux du Napping® est de fournir un cadre à des sujets leur permettant de s'exprimer librement par rapport à la perception qu'ils peuvent avoir d'un ensemble de stimuli.

Depuis, cette méthode a été maintes et maintes fois éprouvée, parfois dans des contextes qui laissent penser qu'elle a été mal comprise, notamment quant à la nature des données qu'elle fournit.

Un des objectifs de cette thèse est d'apporter un éclairage nouveau allant au-delà des données fournies à l'étape ultime du Napping®, et prenant en compte le comportement du sujet lors de la tâche de Napping®. Cette prise en compte du comportement permet une meilleure compréhension des données issues du Napping® d'une part, une analyse plus fine de ces mêmes données d'autre part.

Ainsi de façon progressive, ce manuscrit nous amène des données de Napping® telles qu'elles ont été originellement imaginées, recueillies et analysées, à des données plus complexes qui prennent en compte le consommateur confronté à la tâche.

La Figure 0.1 est une illustration de la façon dont les données de Napping® originelles évoluent au cours de ce manuscrit. C'est en quelque sorte une illustration du fil conducteur de ce travail de recherche. Le chapitre 1 explique comment, sur la base de l'observation du sujet réalisant la tâche, on est passé de l'analyse des coordonnées des stimuli à l'analyse conjointe de ces mêmes coordonnées et des groupes auxquelles appartiennent les stimuli (cf. Fig. 0.1a). Cette information supplémentaire issue du comportement du sujet, amène naturellement au chapitre 2, dans lequel on se propose de modéliser ce comportement à partir de la configuration finale. Ce modèle nous permet d'obtenir une information supplémentaire essentielle à la compréhension des données, puisqu'il s'agit des différentes étapes par lesquelles passe le sujet lors de la réalisation de la tâche (cf. Fig. 0.1b). On introduit dans ce chapitre une notion importante, celle du processus cognitif du sujet qui correspond à l'ensemble des étapes qui permettent d'aboutir à la configuration finale fournie par le sujet, autrement dit aux données de Napping® originelles. C'est finalement cette notion importante de processus cognitif que l'on étudie plus en détail dans le chapitre 3. Dans ce chapitre, on propose une nouvelle méthode qui permet le recueil de ces données. Cette méthode, appelée digit-tracking, permet d'enregistrer l'ensemble des étapes qui conduisent à la configuration finale du sujet (cf. Fig. 0.1c).

En résumé, ce travail de recherche expose comment les données de Napping® peuvent s'enrichir progressivement en prenant en compte une information de plus en plus fine sur le comportement du sujet, à différents niveaux. Dans le chapitre 1, cette information se traduit par l'observation d'une variable qualitative d'appartenance à un groupe, autrement dit une partition. Dans le chapitre 2, cette information se traduit par une suite de partitions emboîtées. Dans le chapitre 3, cette information correspond à l'observation en temps réel du processus cognitif du sujet, où l'on s'affranchit finalement de la contrainte d'emboîtement sur les partitions. Enfin dans le chapitre 4, nous présentons un environnement de travail, appelé Holos, consacré au recueil de données de Napping® temporelles d'une part et à leur analyse d'autre part. Cette analyse permet notamment d'obtenir une représentation du processus cognitif d'un sujet au cours du temps.

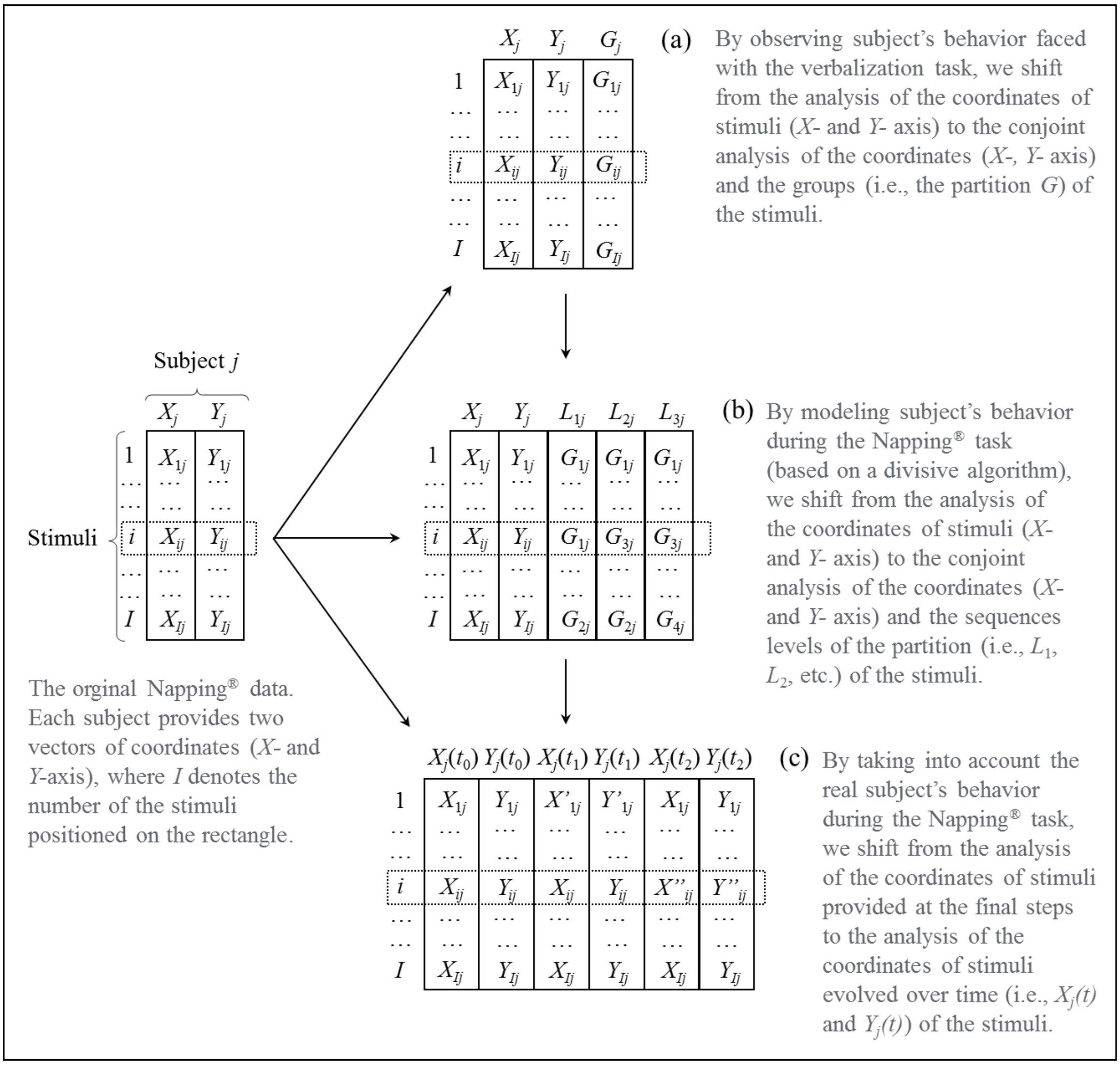


Figure 0.1 Évolution de données de Napping® au cours du manuscrit

Introduction

L'analyse sensorielle a connu une révolution au cours de ces quinze dernières années, à la fois dans sa pratique et dans son utilisation. Les contraintes inhérentes à l'obtention d'un profil sensoriel par exemple, à savoir des coûts et un laps de temps élevés, combinées à une évolution du rapport qu'ont les individus au temps, au changement et à la durée de vie des biens de consommation, ont amené les industriels et les chercheurs à réfléchir à de nouvelles stratégies de développement produits, pour à la fois aller plus vite et au plus près du consommateur.

Une frontière a été franchie lorsqu'il s'est agi de faire participer des sujets non entraînés au processus d'évaluation sensorielle, jusqu'alors réservé aux seuls sujets entraînés. Dans un premier temps, il a été montré que l'on pouvait obtenir des profils sensoriels comparables en partant de deux instruments de mesure de nature différente, un panel entraîné et un panel naïf (Husson, Le Dien, & Pagès, 2001). Dans un deuxième temps, le panel naïf, souvent composé de consommateurs, a été utilisé pour lui-même, non pas pour sa capacité à mesurer, mais pour ce qu'il mesurait : le point de vue du consommateur, qu'il fallait intégrer dans la développement de nouveaux produits (Costa & Jongen, 2006; MacFie, 2007; Whitley, 2010).

Cette intégration du consommateur dans le processus de recherche et développement de nouveaux produits a orienté la recherche vers de nouvelles méthodes de recueil de données. Une des directions prises à travers ces orientations, est celle qui consiste à utiliser des méthodes moins analytiques que celle du profil sensoriel (e.g., Quantitative Descriptive Analysis, QDA®; Stone et al., 1974). Ces approches sont parfois qualifiées d'holistiques, puisqu'elles considèrent l'objet d'étude comme un tout et non pas comme une somme de composantes ou de caractéristiques, par opposition au profil sensoriel où par définition un profil est un ensemble d'éléments caractéristiques.

Parmi ces approches, on peut citer le tri et ses différentes déclinaisons (Chollet et al., 2011), ainsi que le Napping® (Pagès, 2005), à la fois proche conceptuellement du Projective Mapping (Risvik, McEwan, Colwill, Rogersa, & Lyonb, 1994; Risvik, McEwan, & Rødbotten, 1997), mais néanmoins très différent à quelques égards.

De nombreux travaux de recherche ont été consacrés au tri et ses déclinaisons et applications dans le domaine de l'analyse sensorielle. Au cours de ces 5 dernières années, deux thèses notamment ont été entièrement consacrées à l'analyse des données de tri (Cadoret, 2010; Faye, 2012).

Peu a été écrit ou dit sur la tâche de Napping® elle-même, alors que cette dernière connaît un engouement qui ne semble pas faiblir, en témoignent les articles (Giacalone, Bredie, & Frøst, 2013; Kim, Jombart, Valentin, & Kim, 2013), chapitres de livre (Dehlholm, 2014; Lê, Lê, & Cadoret, 2015), et logiciels (TimeSens, ChemoSens; Fizz, Biosystemes; Logic8, et EyeQuestion) parus ces 5 dernières années.

En revanche, cette tâche a été de nombreuses fois comparée à d'autres protocoles de recueil de données sensorielles, sur la seule base des espaces produits obtenus par l'une ou l'autre des méthodes de recueil. Cette comparaison est de notre point de vue déplacée dans la mesure où le Napping® et le QDA®, méthode de référence à laquelle le Napping® est souvent comparé, n'ont pas les mêmes objectifs. Dans les paragraphes suivants, nous rappelons brièvement le protocole du Napping®, puis son objectif principal.

Dans le Napping®, il est demandé au sujet de positionner des stimuli sur une nappe selon leurs proximités perçues. Autrement dit, deux stimuli vont être d'autant plus proches physiquement qu'ils ont été perçus comme proches. Cette nappe a pour particularité d'être sous forme d'un rectangle dont la longueur est 1.5 fois plus grande que la hauteur. Ce sont précisément ces détails qui permettent de différencier le Napping® du Projective Mapping (PM, Risvik et al., 1994) et qui confèrent aux données recueillies par Napping® des règles d'interprétation spécifique. Ce point de vue est partagé par certains auteurs, comme par exemple Dehlholm et al. (2012), Frøst et al. (2015), qui distinguent ces deux méthodes de façon rigoureuse. Il ne l'est pas par d'autres, comme par exemple Nestrud & Lawless (2010) ou Hopfer & Heymann (2013), qui utilisent de manière interchangeable les deux termes PM et Napping®.

L'objectif déclaré du Napping® est de faire émerger les principales dimensions de variabilité perçues par un consommateur à travers un ensemble de stimuli, et de les hiérarchiser ; ces dimensions étant propres à chaque consommateur. Cet objectif ne ressemble en rien à celui du QDA® qui se base sur une liste d'attributs et donc un point de vue déterminé a priori pour fournir une description des stimuli en fonctions de leur profil sensoriel.

A l'instar de la thèse de Marine Cadoret sur le traitement statistique de la tâche de tri et ses déclinaisons, les objectifs de ce travail de recherche sont d'apporter des éléments de réflexion autour du Napping® et des données que cette tâche fournit. Ces éléments de réflexion portent essentiellement sur l'intégration progressive du comportement du sujet dans l'analyse des données de Napping® telles qu'elles ont été imaginées originellement.

Originellement, les données de Napping® avaient pour seule fin la mesure de la distance inter-stimuli perçue par un sujet. On ne s'intéressait donc, presque exclusivement, qu'aux seuls produits, et non pas à la personne qui les mesurait. Or comme le fait justement remarquer Delarue (2015, p.5): *"it's now widely assumed that sensory measurements are subjective, and essentially capture a product × subject interaction*". Ce concept d'interaction est d'autant plus important aujourd'hui que le sujet est un consommateur : la compréhension des données restreintes aux seuls produits ne suffit plus, il faut y intégrer des informations relatives aux sujets prises au cours de la tâche.

Le chapitre 1 de ce manuscrit s'attache à retracer l'évolution du Napping® en Napping catégorisé. Cette évolution correspond à la première prise en compte du comportement du sujet dans les données de Napping®. Les données initialement limitées aux coordonnées des stimuli sur la nappe de dimension 60 par 40 cm se voient complétées par une variable qualitative, qui correspond à l'appartenance des stimuli à un groupe, en fonction de la façon dont ils ont été regroupés puis décrits. On ne s'intéresse pour l'instant qu'aux données recueillies à la toute fin de la tâche, mais on a le sentiment qu'une prise en compte d'informations en amont permettrait une meilleure compréhension de l'espace produit, de l'interaction *product* × *subject*.

Le chapitre 2 de ce manuscrit propose dans un premier temps un modèle du comportement du sujet lors de la tâche de Napping®, dans un second temps l'intégration de ce modèle comme une information supplémentaire permettant de mieux comprendre l'espace produit. C'est au cours de ce chapitre 2 qu'apparaît la notion centrale de processus cognitif, que l'on peut définir abusivement dans le contexte de ce manuscrit comme la suite des étapes qui amènent un sujet à sa configuration finale. Le modèle proposé, bien qu'étant intéressant, ne prend pas en compte les changements éventuels de perception qui peuvent se produire au cours de la tâche.

Le chapitre 3 présente le cœur de notre recherche, à savoir la prise en compte du processus cognitif observé dans l'analyse des données de Napping®. Cette prise en compte nécessite un recueil de données et une analyse de ces données ad hoc. On introduit dans ce chapitre le concept de digit-tracking qui mesure au cours du temps la proximité entre les stimuli. Ce nouveau type de données temporelles a nécessité le développement de nouveaux outils, de recueil et d'analyse, que nous présentons dans le chapitre suivant.

C'est dans le chapitre 4 que nous expliquons comment la plateforme de recueil de données Holos a été conçue. Cette plateforme couplée à une application Android installée sur une tablette tactile, permet de recueillir des données de Napping® qui prennent en compte le processus cognitif du sujet, du début à la fin de la tâche. Holos a été pensé comme un environnement collaboratif qui permet au chercheur de partager ses stimuli, voire ses données. Dans ce même chapitre, nous expliquons comment les données de Napping® peuvent être analysées avec les nouvelles fonctions du package SensoMineR développées à cet effet.

Les 3 premiers chapitres de ce manuscrit sont principalement constitués d'un chapitre d'ouvrage paru en 2015, d'un article dans une revue à comité de lecture paru en 2015, et d'un article en cours de soumission (2014). Ils sont également constitués d'éléments de réflexion complémentaires situés à la fin des chapitres.

(This page is left blank)

1

Napping® and sorted Napping as a sensory profiling technique

Integrating subject's behaviour issued from verbalisation task in analysing Napping® data

Si l'on se réfère à la littérature, il semblerait que la caractéristique première retenue par les utilisateurs du Napping® soit sa rapidité d'utilisation : dans la littérature anglaise, on parle de *rapid sensory profiling method*.

Contrairement aux méthodes descriptives classiques (e.g., Quantitative Descriptive Analysis, QDA®; Stone et al., 1974, or Spectrum®, Munoz & Civille, 1992), la réalisation d'une tâche de Napping® ne nécessite pas d'entraînement particulier, pour peu que le sujet qui la réalise ait au minimum la capacité de se repérer sur un plan. Cette caractéristique essentielle pour les utilisateurs du Napping®, dans un contexte où tout va toujours plus vite, a occulté ce pour quoi cette méthode a été créée, à savoir sa capacité à fournir un point de vue issu de consommateurs. De façon restrictive, pour de nombreux utilisateurs, le Napping® est un moyen d'obtenir relativement rapidement une représentation graphique d'un ensemble de produits dans un espace à deux dimensions. Vu sous cet angle, on peut comprendre qu'il soit légitime, à première vue d'évaluer le Napping® en le comparant à la méthode de référence jusqu'alors utilisée pour obtenir une représentation graphique d'un ensemble de produits (Risvik et al., 1997; Kennedy & Heymann, 2009; Dehlholm, Brockhoff, Meinert, Aaslyng, & Bredie, 2012; Louw et al., 2013; Cadena et al., 2014) : existe-t-il des différences entre les représentations d'un espace produit fournies par un profil sensoriel classique d'une part, par le Napping® d'autre part ?

Cette comparaison, selon nous, ne permet pas d'arbitrer entre Napping® et profil sensoriel classique, tant la nature des données utilisées pour construire les représentations graphiques est différente d'un recueil de données sensorielles à l'autre. Elle témoigne même d'un manque de compréhension des données issues du Napping®.

L'objectif de ce chapitre est d'apporter des précisions sur la nature de ce qui est mesuré lors d'une tâche de Napping®, à savoir une information holistique propre à chaque sujet et non dépendante d'une cadre fixé a priori comme c'est le cas pour le profil sensoriel classique.

La rapidité que certains confèrent au Napping® est essentiellement due à l'absence de séance d'entraînement. Mais, ceux qui ont pratiqué le Napping® savent que la réalisation même de la tâche peut être longue pour un sujet donné pour qui la tâche peut s'avérer complexe. Une meilleure compréhension de ce qui est mis en œuvre par un sujet lors de la tâche de Napping® semble être un élément important pour comprendre sa représentation de l'espace produit. Le deuxième objectif de ce chapitre est de présenter les prémices de nos travaux de recherche sur la prise en compte du comportement du sujet lors de la tâche de Napping®.

Ce chapitre est constitué d'un chapitre de livre : Napping® and sorted Napping as a sensory profiling technique (Lê, Lê, & Cadoret, 2015) et d'un complément de réflexions qu'il nous a semblé intéressant d'intégrer afin de préciser notre pensée.

Chapter outline

This section includes the publication:

Lê, S., Lê, M.T., & Cadoret, M. (2015). Napping® and sorted Napping as a sensory profiling technique. In J. Delarue, B. Lawlor, & M. Rogeaux (Eds.), Rapid sensory profiling techniques and related methods (pp. 197–213). United Kingdom: Woodhead Publishing. doi: 10.1533/9781782422587.2.197

Abstract

Can (sorted) Napping® be considered a rapid method? Can it be considered as an alternative to descriptive analysis? The main objective of this chapter is to shed new light on Napping® and sorted Napping® in order to provide some clues for answering these two questions. In particular, we will stress the intrinsic nature of the data collected when using Napping®, and on the link between Napping® and sorted Napping. We will see how the information provided by Napping® is unique, and how different it can be from that provided by descriptive analysis. We will see that, at a subject level, Napping® is certainly a rapid method, but in order to be used as an alternative to descriptive analysis, a relatively high number of subjects need to be considered.

Keywords: Projective test, Napping®, Verbalization, Sorted Napping, Sensory profile.

1.1 Introduction

This chapter is divided into two main parts. The first is dedicated to Napping®, while the second is dedicated to sorted Napping. We will see finally how close Napping® and sorted Napping are, and how they can easily be confused in a large number of situations. Each part is illustrated by an example.

The first example used to illustrate Napping® was chosen for two main reasons. First, it will allow the reader to understand the intrinsic nature of the data collected when performing Napping®, and the natural evolution from Napping® to sorted Napping. Second, although the example is not directly related to sensory profiling, it is typical of the sensations sensory researchers are confronted with when they develop a new product, as it deals with emotions.

The second example is directly focused on the measurement of sensory perceptions of food products. It is a pedagogical example that illustrates perfectly well how sorted Napping can be used as a sensory profiling technique. In this example, eight smoothies were chosen according to two main factors: the type of manufacturer (two categories) and the flavour (four categories).

1.2 From projective test to Napping®

1.2.1 An introduction to projective test

In psychology, notably in clinical psychology, projective tests were conceived to assess the personality of a patient. Such tests allow subjects to indirectly reveal their personalities by experiencing different, vague, ambiguous stimuli, and by projecting themselves through their responses to the stimuli, hence the adjective projective. One of the most famous and controversial projective tests is certainly the Rorschach inkblot test.

This family of tests has been adapted for, and then frequently used in, marketing research. We can cite for instance the word association task in which "the subjects are asked to read a list of words and to indicate the first word that comes to mind." For a short and easily available review of projective tests in consumer research, refer to Donoghue (2000).

In the sensory field, in 1994, Risvik *et al.* published a paper entitled "Projective Mapping: a tool for sensory analysis and consumer research": the idea of asking subjects to reveal themselves through the way they are positioning stimuli (products) on a sheet of paper based on their perceived similarities arose in sensory analysis. However, from our point of view, this brilliant idea was not exploited for its proper use. While many different angles could have been approached, the authors focused on a representation of the stimuli obtained by using generalized Procrustes analysis (GPA), and its comparison with other representations issued from sensory profiling, and dissimilarity scaling. This is, somehow, not very surprising, as the reasoning of Risvik et al. (1994) seems to be mainly based on the stimuli and their differences, not on the subjects and the way they perceive differences between stimuli.

From our point view, the idea that "when an assessor evaluates the products as a whole he/she will concentrate on the most obvious differences when making a judgment" requires further discussion. Beyond the differences that may exist between stimuli, it seems that one crucial dimension has been overlooked: the individual variability amongst the subjects. In other words, what may be obvious for one subject may not matter for another? Considering these differences of perception between subjects is particularly important in our context, as the perception of the stimuli is holistic, on the one hand, while the way this perception is expressed is projective, on the other hand. As explained in the next section, this is the starting point of Napping®.

1.2.2 The intrinsic nature of Napping® data

Napping® arrived in 2005, in "Collection and analysis of perceived product inter-distances using multiple factor analysis (MFA): Application to the study of ten white wines from the Loire Valley" (Pagès, 2005). In its basic version, Napping® consists in asking subjects to position stimuli on a "nappe" (i.e., French word for tablecloth; originally, the "nappe" was simply a sheet of paper with specific dimensions) according to the way they perceive their similarities. In practice, each subject uses his own criteria to position stimuli. Two stimuli should be closer to the degree they were perceived to be similar, and more distant as they were perceived to be different. There are no good or bad answers.

With respect to the way data are collected, the two methods, Projective Mapping and Napping®, look very much alike. However, the context in which each method was conceived, and their respective motivations, are very much different. Napping® was developed in a period where consumers began to be used to providing sensory data. At that time, the paradigm that would strictly separate sensory data provided exclusively by trained panelists, and hedonic data provided by consumers, was evolving and consumers were more and more solicited for providing sensory information. Not any kind of sensory information, as you do not expect consumers to be as accurate as trained panellists on a list of sensory descriptors, but a more qualitative one that would give insights about what consumers can sensorily perceive.

As mentioned by Pagès in his article, when using a conventional sensory profile based on a given list of sensory descriptors, "one criticism of this methodology is that the weights given to the descriptors do not necessarily correspond to their real importance for the subjects". As evoked previously, this was the main starting point of Napping®: to find a way that would reveal what is most important (or obvious) for subjects, in particular consumers, in terms of perceived differences amongst stimuli. How to get this information without asking? The answer lies in the dimensions of the "nappe".

The dimensions of the "nappe," its width and its height, are of utmost importance. The "nappe" has a rectangular shape, with a width of 60 cm and a length of 40 cm. Consequently, as subjects are naturally influenced by the X-axis, which is by construction 1.5 visually more important than the Y-axis of the "nappe," subjects are prioritizing the reasons why they perceive the stimuli differently. More importantly, this hierarchy is obtained unconsciously and spontaneously. In other words, the rectangle in which stimuli are positioned leads subjects to privilege one dimension (i.e., the X-axis) and to separate stimuli with respect to that dimension, then to use the second dimension (i.e., the Y-axis), in a hierarchical order. The idea of Napping® is, first, to collect, through the positioning of the stimuli, the first and the second reason why stimuli are perceived as different, for each subject; then, to summarize all these reasons over all the subjects. To do so, Napping® data need to be analysed specifically, with the proper method. In that sense, Napping® can be seen as the combination of the way data are collected and analysed, as will be explained in the next section.

1.2.3 Analysing Napping® data

As a result of Napping®, each subject provides two vectors of coordinates of dimension $I \times 1$ each (one for the *X*-axis, one for the *Y*-axis), where *I* denotes the number of stimuli to be positioned on the rectangle. Hence, the final data set to be analysed, denoted *X*, is obtained by merging the *N* pairs of vectors of coordinates, where *N* denotes the number of subjects. In other words, *X* can be seen as a data set structured into *N* groups of two variables each. Typically, the statistical analysis of such data set *X* should take into account the "natural" partition on the variables.

MFA (Escofier & Pagès, 1994; Pagès, 2004, 2013) was precisely conceived in order to balance the part of each group of variables within a global analysis. When variables are quantitative (which is our case), MFA can be defined as an extension of principal component analysis (PCA), which takes into account a group structure on the variables, and which balances the part of each group: in that sense, MFA can be seen as a weighted PCA. Technically, each group *j* of variables is weighted by the inverse of λ_1^j, where λ_1^j denotes the highest eigenvalue issued from the PCA performed on the variables of group *j*.

The rationale behind this weighting scheme is to exhibit dimensions that are: (1) common to as many groups of variables as possible; and (2) specific to some groups of variables. From this perspective, MFA can be defined as a variant of Generalized Canonical Analysis (Carroll, 1968).

When analysing Napping® data, in order to respect the intrinsic nature of the data, λ_1^j is obtained by running a PCA on the covariance matrix, not on the correlation matrix. In other words, the separate analyses of each group of coordinates are not performed on standardized variables: the *X*-coordinates and the *Y*-coordinates of each group are not scaled to unit variance. Otherwise, the information about the relative importance of the dimensions of variability amongst the stimuli would be removed: this information is precisely what the experimenter is looking for when he/she is collecting data using Napping®.

1.2.4 Example: revealing oneself with Napping®, through emotions as stimuli

The data presented in this section are part of a cross-cultural study of which the main purpose is to compare two very supposedly different countries, Vietnam and France, through their respective emotional spaces (i.e., the way emotions are related to each other). For this comparison, we asked Vietnamese and French subjects to perform a Napping® on a set of 53 emotions. The emotions were chosen in order to obtain a corpus of words that would be as diverse and comprehensible as possible. Pre-tests were conducted to make sure that the corpus was easily understandable. This study can be divided into two parts, corresponding to two main objectives, namely (1) to understand how emotions were perceived and structured within each country, and (2) to compare Vietnam and France in terms of emotions. To illustrate how Napping® can be used to reveal oneself, through emotions as stimuli, we present here results issued from the first part of the study, restricted to the French data (100 French subjects).

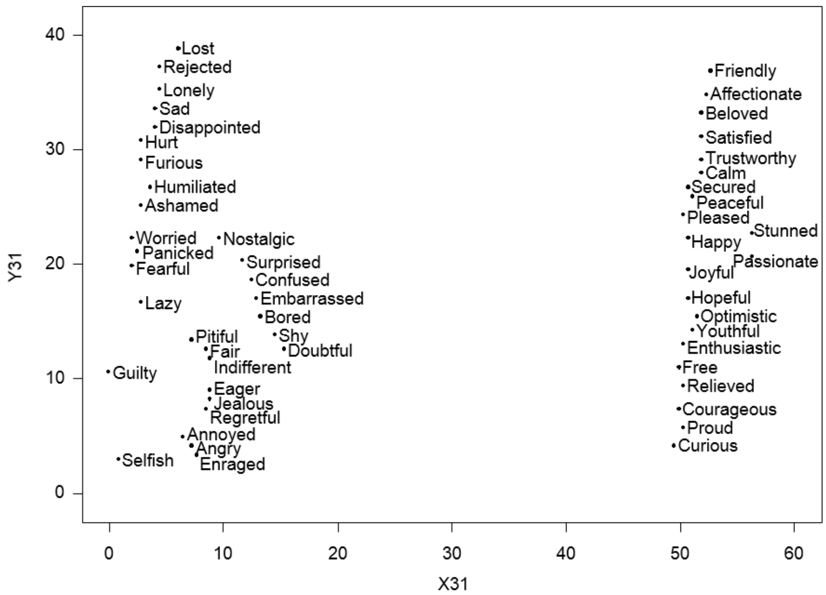


Figure 1.1 Representation of the emotions provided by Subject 31

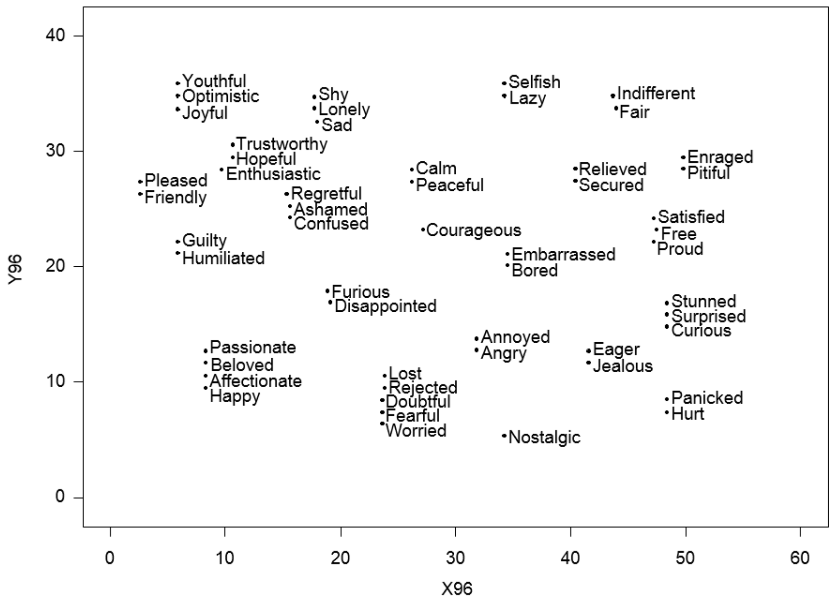


Figure 1.2 Representation of the emotions provided by subject 96

Figure 1.1 and Figure 1.2 show two completely different individual configurations. The configuration of Subject 31 (cf. Figure 1.1) is very structured: emotions seem to have been divided into two groups, along the *X*-axis, the positive emotions (on the right-hand side of the plane) and the negative ones (on the left-hand side of the plane). The configuration of Subject 96 (cf. Figure 1.2) seems to be structured as well, but in a different way. Subject 96 provided small clusters of emotions, but the positioning of the clusters seems to be hardly interpretable. Subject 96 may not have understood the Napping® task, and hence comes the question of getting a representation of the stimuli based on all subjects.

To answer that question, an MFA was performed on the whole of the data. In detail, MFA is performed on groups of two variables each (i.e., the *X*- and *Y*-coordinates of the stimuli on each subject's rectangle). Let us remind ourselves that the separate analysis of each group is based on unstandardized variables.

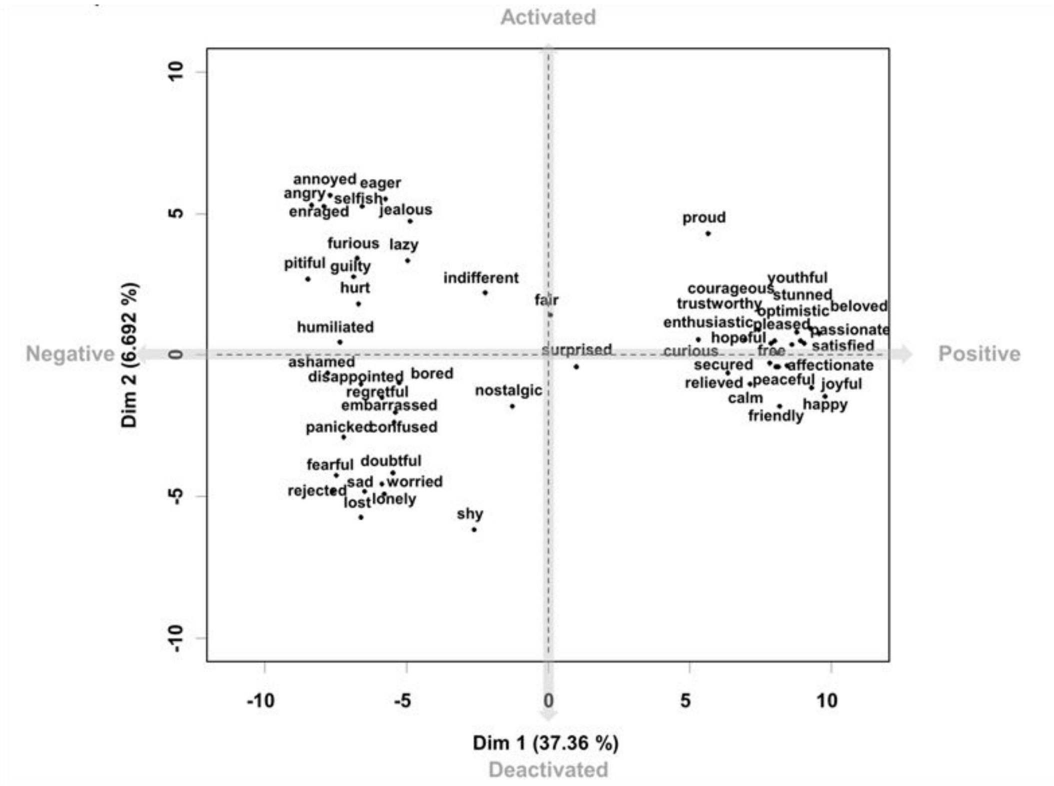


Figure 1.3 Representation of the emotions issued from MFA

Figure 1.3 corresponds to the representation of the 53 emotions (considered as statistical units) obtained by using MFA. This representation is the consensus over the subjects on the so-called first factorial plane, constituted by the first two main dimensions of variability. The first dimension (i.e., the horizontal one), denoted Dim 1 in Figure 1.3, explains 37.36% of the total variance. This relatively high value (considering the number of individuals and variables in the analysis) indicates that there is a consensus amongst subjects.

This first dimension clearly opposes positive emotions such as "happy" and "joyful" (on the right-hand side of the plane) to rather negative emotions such as "lost" and "sad" (on the left-hand side of the plane).

The second dimension, denoted Dim 2 in Figure 1.3, explains 6.69% of the total variance. Although this value seems apparently low, this dimension will be kept in the analysis since it has been tested as statistically significant by using bootstrap technique (not presented in this chapter), and is obviously interesting in terms of interpretation.

Indeed, amongst the so-called negative emotions, this second dimension opposes emotions such as “angry” and “furious,” which could be described as activated, to emotions such as “worried” and “lost,” which could be described as deactivated.

Let us notice the singular position of the emotion “surprised” located almost on the centre of gravity of the plane. It seems that it has been isolated by the subjects, and that it has been perceived as neither positive nor negative: one can be “surprised” in positive and negative ways; “surprised” could easily be qualified as “neutral.”

In terms of dispersion, we can notice the difference between the group of so-called positive emotions and the group of so-called negative emotions: positive emotions seem to be more homogeneous than negative emotions, which are more heterogeneous. Positive emotions are first perceived as a whole, which may be due to their strong common character (i.e., positiveness) perceived by the subjects; whereas negative emotions are perceived more as part of some kind of continuum. Actually, the third dimension (not shown here) also represents the positive emotions more as a continuum.

To supplement the previous findings, Table 1.1 shows that the first eigenvalue associated with Dim 1 equals 49.43, for a maximum value of 100 if all hundred subjects had expressed exactly the same first dimension (due to the weighting, the maximum value of the highest eigenvalue in MFA equals the number of active groups considered in the analysis). This value is indicative of the fact that subjects share a common important dimension that opposes positive emotions to negative ones.

Table 1.1 Decomposition of the total variance on the first 10 dimensions

	Eigenvalue	Percentage of variance	Cumulative percentage of variance
Dim 1	49.43	37.36	37.36
Dim 2	8.85	6.69	44.05
Dim 3	5.58	4.22	48.27
Dim 4	4.85	3.67	51.94
Dim 5	4.03	3.05	54.98
Dim 6	3.79	2.86	57.85
Dim 7	3.66	2.76	60.61
Dim 8	3.36	2.54	63.15
Dim 9	3.23	2.44	65.59
Dim 10	2.71	2.05	67.64

The consensual representation in Fig. 1.3 can also be understood regarding the way each subject has positioned the emotions during the experiment. To do so, the results of the separate analyses of each group of coordinates (one group corresponding to one subject) can be projected as supplementary information on the axes provided by MFA.

In the case of our two Subjects 31 and 96 (cf. Figs 1.1 and 1.2), Figure 1.4 shows that the main axis of variability of Subject 31 (denoted Dim 1.S31) is almost perfectly correlated to the main axis of variability of MFA based on all subjects. It is also true, to a lesser degree, for the second dimension of Subject 31 (denoted Dim2.S31), whereas it is completely false for the dimensions of Subject 96 (Dim 1.S96 and Dim 2.S96).

Complementary to that output, MFA provides a representation of the groups of variables, the subjects in a Napping® context (cf. Figure 1.5). This representation is based on the axes provided by MFA.

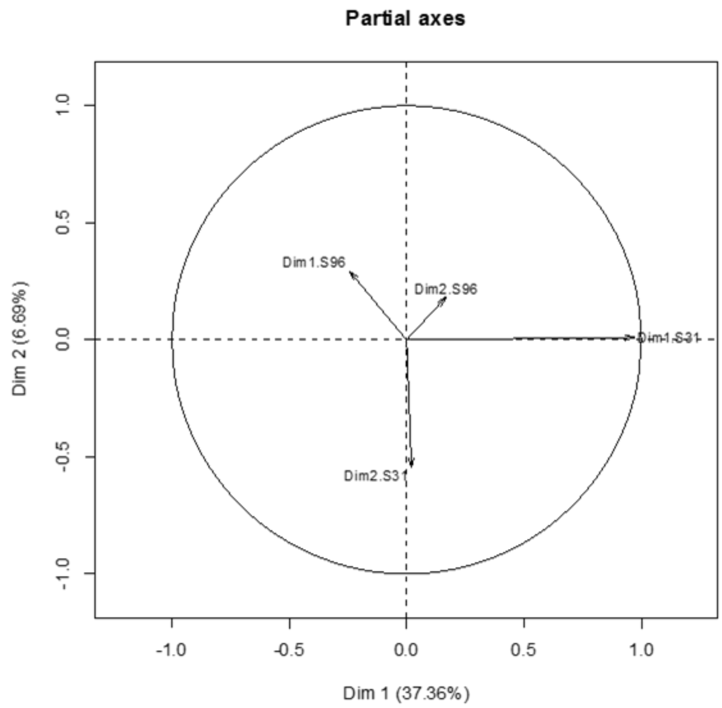


Figure 1.4 Representation of the axes issued from the separate analyses of the groups associated with subjects 31 and 96

Beyond the optimal properties of this representation that will not be detailed here, coordinates of a subject on an axis lie between 0 and 1, and can be interpreted as the importance of the axis given by the subject with respect to his own stimuli configuration. When this value equals 0, as is almost the case for Subject 96 on the first axis, it means that the structure induced by the axis provided by MFA has nothing to do with either of the two coordinates provided by the subject. On the other hand, when this value equals 1, as is almost the case for Subject 31 on the first axis, it means that the structure induced by the axis provided by MFA can be assimilated to the principal dimension of variability of the subject. The opposition between positive and negative emotions on the first axis of MFA corresponds to the main axis of variability for Subject 31, while it does not mean anything for Subject 96. According to Figure 1.5, it seems that, for a majority of subjects, the opposition between positive emotions and negatives ones represents an important dimension of variability.

This example is remarkable in the sense that dimensions are directly interpretable in terms of stimuli. This is mainly due to the nature of the stimuli, which have a clear definition and can be interpreted at will. By analogy, a representation of a product space is hardly interpretable without expert knowledge about the products; in the case of a sensory profile, without the representation of the sensory descriptors.

Supplementary information appears to be essential for interpreting the relative positioning of the stimuli (usually the products), as this sole information is generally not self-contained. Hence came the idea to supplement the original information from Napping® (i.e., the only coordinates of the stimuli, as in the example), by asking subjects to add comments on the stimuli, once the placement of the stimuli had been done: this is what has been called verbalization, literally to express thoughts, feelings and emotions in words.

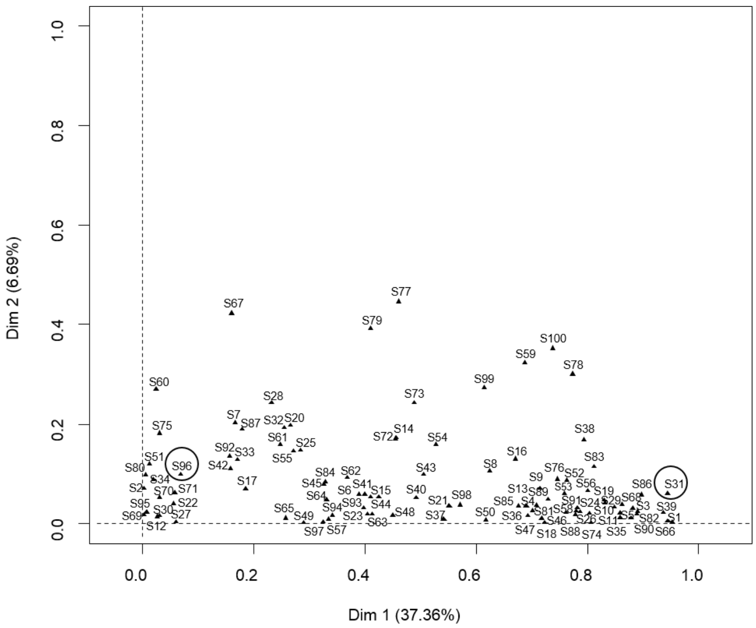


Figure 1.5 Representation of the groups of coordinates in MFA (i.e., subjects in a Napping® context)

Originally, this information was aggregated over the subjects and coded into a contingency matrix (stimuli × words), where rows correspond to stimuli, columns correspond to words, and at the intersection of one row and one column, the number of times that the word was used to qualify the stimulus. Then, this matrix was usually considered in MFA as a supplementary group. In other words, this matrix was added to the matrix of coordinates, then treated as a supplementary group, while the groups of coordinates were treated as active groups. The information related to the contingency table was represented on a correlation circle, as if each column was assimilated to a continuous variable.

The representations of the stimuli and of the verbalization data provided by the subjects are actually very similar (including in terms of usage) to the representations issued from classical sensory profile data analysed by PCA. From this perspective, Napping® combined with verbalization can be seen as a sensory profiling technique (as, indeed, Napping® provides a sensory profile).

Still, intrinsically, the information provided by Napping® is unique: this information is based on a multitude of sources of differences amongst the stimuli, these sources being important for the subjects. Without wanting to sound polemical, Napping® is neither a rapid method, nor an alternative to descriptive analysis. More precisely, at a subject level, Napping® is intrinsically a rapid method, as each subject is considering the stimuli globally, and as we expect from each subject a spontaneous answer corresponding to her/his most striking differences amongst the stimuli. But to fully exploit this unique information, the number of subjects usually involved is rather high, which extends the length of the experiment. The uniqueness of the information provided by the subjects makes us believe that Napping® is not *sensu stricto* an alternative to descriptive analysis: we do not expect the representations of the stimuli from a classical descriptive analysis and from a Napping® to be similar; the first is based on a fixed list of descriptors, while the second is based on a multitude of sources of differences amongst stimuli, unknown a priori.

Now that we have understood that verbalization is of utmost importance, and that it should be systematically asked of subjects to interpret the results, we will see in the next section where sorted Napping originated.

1.3 From Napping® to sorted Napping

1.3.1 Some behavioural considerations

The idea of sorted Napping arose from the observation of subjects during the verbalization phase (Cadoret et al., 2009; Pagès et al., 2010). In effect, it appeared that most of the subjects were not verbalizing stimulus by stimulus. Instead of an individual description of the stimuli, we observed that subjects tended to provide a description of the stimuli at a more global level, with respect to clusters of stimuli. In practice, when they were asked to verbalize, subjects were naturally making clusters of stimuli by circling them with their pen, and were describing the stimuli at a cluster level. We named this natural response pattern "sorted Napping" (derived from the sorting task).

This behavioural pattern may have several explanations. One of them could be that the intrinsic nature of Napping® is holistic, and that is precisely what is expected. In a sensory profile context, Napping® will naturally reveal the main sensory dimensions of variability amongst the stimuli (the ones perceived by the subjects). But when it comes to describing the stimuli, subjects have to switch to a more analytical way of thinking: putting words on something that was latent and that has just been revealed may be difficult for the subjects. Hence, the recurrent behavioural pattern that has been observed, and that has shown that subjects were more likely to verbalize at a cluster of stimuli level.

Ultimately, sorted Napping can be assimilated to Napping® followed by a verbalization phase of elicited clusters.

1.3.2 Analysing sorted Napping data

One of the main drawbacks when using a contingency matrix (stimuli × words) to analyse verbalization data is that data are aggregated over subjects. We thus lose a lot of information, as the individual variability amongst the subjects is lost. As a consequence, the other main disadvantage is that the information given by a subject when she/he clusters stimuli in order to describe them is not reported properly. Basically, this information is absolutely meaningful, as it reinforces the notion of distance between stimuli based on Napping® only. In other words, the analyst has at his disposal two types of information, a continuous one based on the distance provided by Napping®, and a categorical one based on the groups provided when circling or sorting the stimuli when verbalizing. In that sense, each subject provides a couple of vectors of coordinates for the stimuli, and eventually a categorical variable that reports the way the subject has gathered the stimuli and has described them.

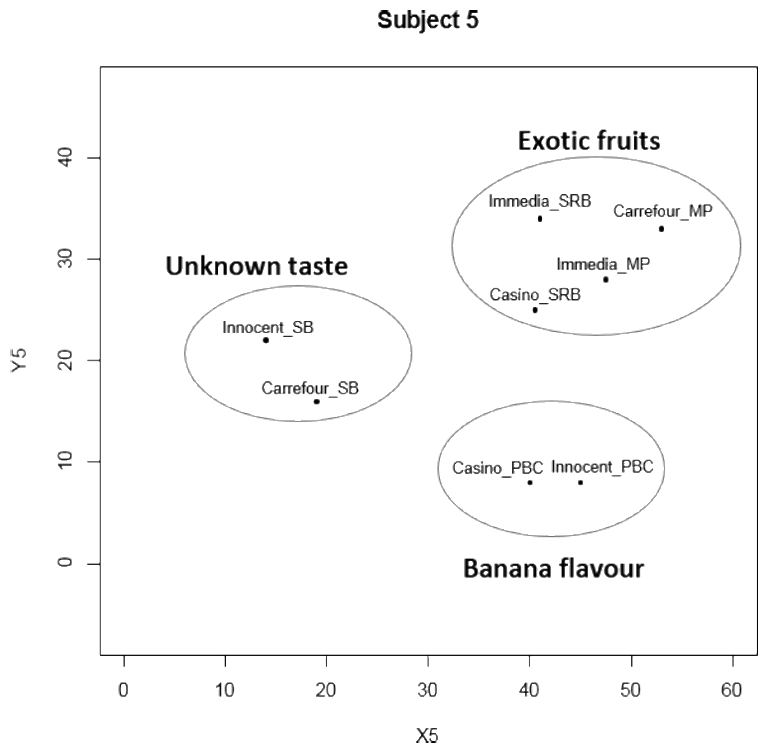


Figure 1.6 An example of data collected during sorted Napping

Table 1.2 Data provided by subject 5, the coordinates and the verbalization data as groups of stimuli

	X5	Y5	C5
Immedia_MP	47.5	28	exotic fruits
Carrefour_MP	53.0	33	exotic fruits
Immedia_SRB	41.0	34	exotic fruits
Casino_SRB	40.5	25	exotic fruits
Innocent_PBC	45.0	8	banana flavour
Casino_PBC	40.0	8	banana flavour
Innocent_SB	14.0	22	unknown taste
Carrefour_SB	19.0	16	unknown taste

The example used to illustrate sorted Napping is based on an experiment in which eight smoothies were chosen according to two main factors, the type of manufacturer (two categories: generic products from retailers such as Casino and Carrefour, and products from national brand name manufacturers such as Innocent and Immedia) and the flavour (four categories: Strawberry-Raspberry- Blueberry denoted "SRB," Pineapple-Banana-Coco denoted "PBC," Strawberry-Banana denoted "SB" and Mango-Passion fruits denoted "MP"). Twenty-four subjects were involved in the experiment. The experiment was conducted by a group of Master students from Agrocampus Ouest, France, who collected the data and who exploited them to win the Syntec Trophy in 2009 (Trophée Syntec des études Marketing & Opinion), one of the most competitive marketing contests in France.

Figure 1.6 represents the rectangle provided by Subject 5, who positioned the eight smoothies first according to the way he/she perceived their similarities, and then sorted them by adding a comment to each group of smoothies. Table 1.2 illustrates the way the information collected in Figure 1.6 is formatted into a data set.

In terms of analysis, it seems rather natural to first try to find a distance amongst the stimuli that combines both continuous and categorical information. Then, once this distance is found, in order to get a global representation, it seems also natural to balance the part of each subject within a global analysis. This is precisely what hierarchical multiple factor analysis (HMFA, Le Dien & Pagès, 2003) does, when considering two levels of partition on the variables:

- A first one, with respect to the subjects, in which each subject is a group of three variables;
- A second one (at the subject level), with respect to Napping® on the one hand, and the sorting on the other hand.

As with Napping®, a block of coordinates is considered an unstandardized continuous group of variables.

1.3.3 Example: (sorted) Napping® as a sensory profiling technique

In this part, we will focus exclusively on the representation of the individuals provided by HMFA, but the graphical outputs and other numerical indicators provided by MFA are provided as well by HMFA, as HMFA is an extension of MFA.

As mentioned previously, the global representation of the individuals (i.e., the smoothies) provided by HMFA is hard to interpret without any supplementary information (cf. Figure 1.7). In this example, the supplementary information the analyst has at his disposal concerns the two factors based on which the smoothies were selected (type of manufacturer and flavour).

In Figure 1.7, we can see that the first factorial plane issued from HMFA differentiates three groups of products. SRB smoothies and Mango-Passion fruit smoothies are opposed to PBC smoothies along the first dimension. Strawberry-Banana smoothies are opposed to the rest of the products along the second dimension. In that sense, we can say that the main structure of variability amongst the smoothies is based on their flavour, not on the type of manufacturer.

Beyond the information based on the factors, the information provided by the subjects on the groups appears to be essential to interpret the factorial plane. This information is represented as categories are represented in Multiple Correspondence Analysis (MCA, Husson, Lê, & Pagès, 2011). In other words, as shown in Figure 1.8, smoothies and description of the groups provided by subjects are represented on the same factorial plane, which is an important and very convenient feature.

Figure 1.8 speaks for itself and demonstrates the capacity of sorted Napping (or Napping® combined with verbalization) to be used as a sensory profiling technique. The comments provided by subjects are rich enough to interpret the dimensions of variability amongst smoothies.

The positive side of the first dimension is characterized by comments such as doughy, sweet, sickly, I don't like, banana, coco, soft, very soft, exotic fruits, creamy texture.

The negative side of the first dimension is characterized by comments such as acid, strong, litchi taste, blueberry, raspberry, with small grains, like guava.

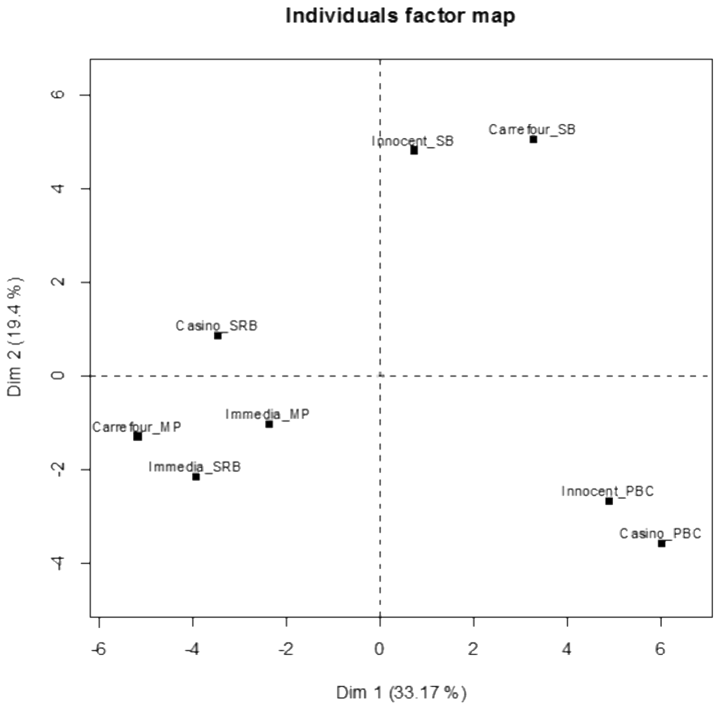


Figure 1.7 Representation of the smoothies based on sorted Napping data

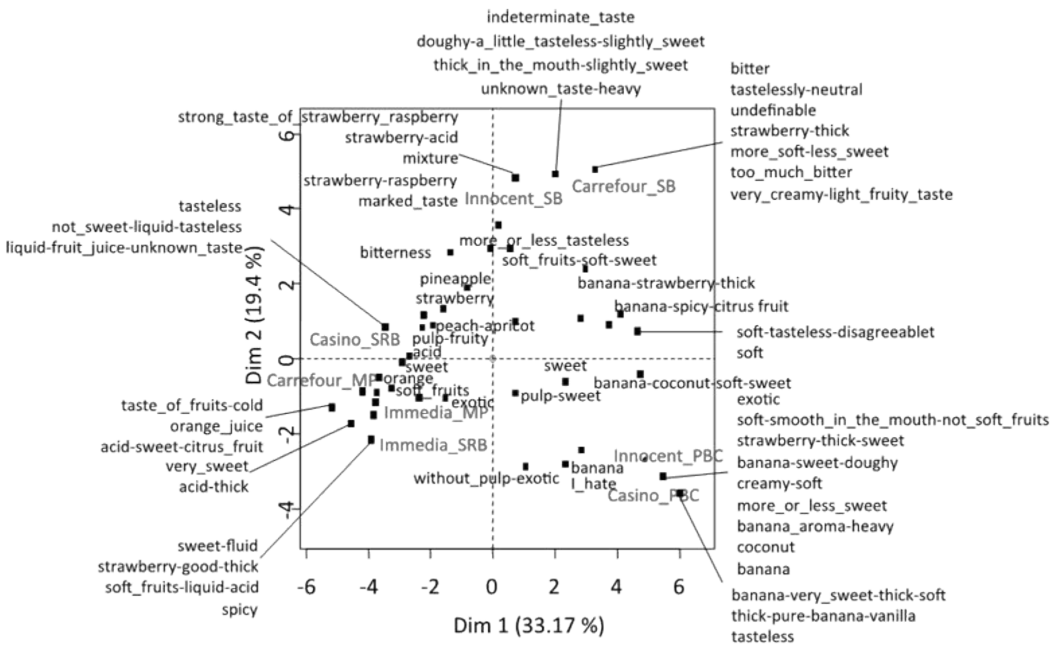


Figure 1.8 Superimposed representation of the smoothies and the comments provided by the subjects

As we can see, smoothies are not all about flavour. An important aspect of this beverage has a lot to do with texture (creamy, soft, thick, liquid, doughy), which depends essentially on the composition of the smoothies.

1.4 Analysing Napping® and sorted Napping data using the R statistical software

The analysis was performed using the statistical software R, the FactoMineR package, an R package dedicated to exploratory multivariate analysis, and the SensoMineR package, an R package dedicated to sensory data analysis.

For Napping® data, the data set comprises 53 rows (statistical individuals) and 200 columns (variables). In other words, 53 emotions and 100 pairs of coordinates, one pair of coordinates corresponding to one subject. In the analysis, each subject is considered as a group of unstandardized variables. Data were analysed using the `MFA()` function of the FactoMineR package. Practically, the code that corresponds to the analysis is the following:

```
> MFA(emotions, group = rep(2,100), type = rep("c",100))
```

The first argument corresponds to the name of the data set, the second corresponds to the way variables are structured into groups and the third corresponds to the way each group is considered in the analysis (as standardized continuous data, as unstandardized continuous data, as categorical data, as a contingency table). The data set emotions on which the analysis has been performed is constituted of a sequence of 100 groups of two variables each (2 has been repeated 100 times), where each group has been considered as a group of unstandardized continuous variables ("c", standing for unstandardized continuous, has been repeated 100 times).

Sorted Napping data were analysed using the `fasnt()` function of the SensoMineR package. Practically, the code that corresponds to the analysis is the following:

```
> fasnt(smoothies, first = "nappe", sep.word = ";")
```

The first argument corresponds to the name of the data set. The second corresponds to the way variables are structured into groups: here Napping® data come first, and sorting data come second. The third argument corresponds to the way comments are coded: ";" between each different comment. This function is really powerful, as it provides a complete description of the products, globally, and individually through the comments (cf. Figure 1.9 and Table 1.3).

The core of the function is based on HMFA. To understand what the function `fasnt()` does, we recommend running an HMFA on the same data. To do so, it is essential to specify the hierarchical structure on the variables, as well as the type of data considered (standardized continuous variables, unstandardized continuous variables and categorical variables). In the case of sorted Napping, the analyst has to specify the fact that subjects provide information based on two unstandardized continuous variables and one categorical variable. When the number of consumers equals 24, for instance, the hierarchical structure is defined in the following way:

```
> list(rep(c(2,1),24), rep(2,24))
```

The first level of the hierarchy is constituted of a sequence of 2 and 1 repeated 24 times (one group of two variables, one group of one variable; repeated 24 times). The second level of the hierarchy is constituted of a sequence of two repeated 24 times (one subject corresponds to the aggregate of two consecutive groups).

Then, the analysis is specified the following way:

```
> HMFA(smoothies, H = list(rep(c(2,1),24), rep(2,24)), type =
        rep(c("c","n"), 24))
```

The first argument corresponds to the name of the data set, the second corresponds to the hierarchical structure on the variables and the third corresponds to the type of data considered in our case a sequence of one group of two unstandardized continuous variables ("c") and one group of one categorical variable ("n" stands for nominal).

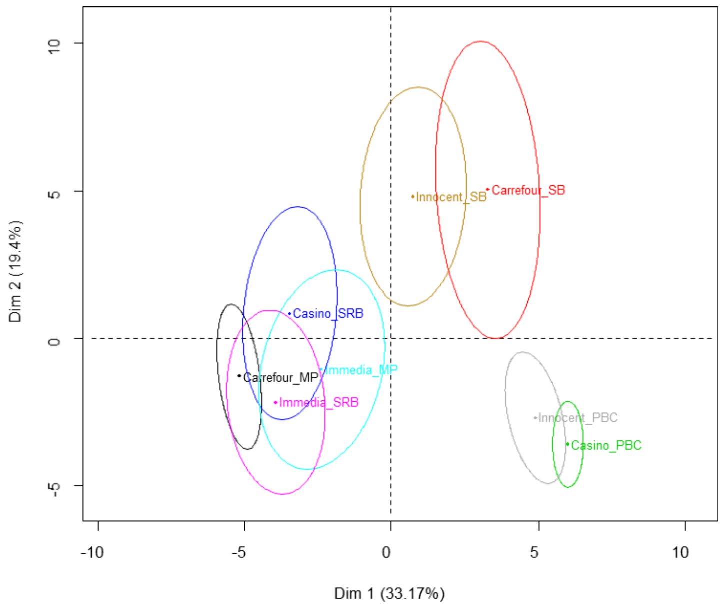


Figure 1.9 An example of graphical representation provided by the `fasnt()` function, representation of the products supplemented by confidence ellipses

Table 1.3 The `fasnt()` function provides an automatic description of the products in terms of verbalization data. Immedia_SRB has been described as acid on the contrary of Innoncent_PBC, the two smoothies being opposite according to the first dimension. Innoncent_SB has been described by the word *strawberry.*

Product	Word	Intern %	Glob %	Intern freq	Glob freq	*p*-value
Immedia_SRB	acid	34.78	15.66	8	26	0.02
Innocent_PBC	acid	0	15.66	0	26	0.01
Innocent_SB	strawberry	22.22	6.62	4	11	0.03

1.5 Conclusion

If used as a sensory profiling technique, Napping® has to be supplemented by a verbalization phase. Otherwise the latent information provided by the subjects through the positions of the stimuli may never be revealed. Observing subjects doing Napping® combined with verbalization led us naturally to sorted Napping and its statistical analysis. In that sense, sorted Napping can be seen as a natural extension of Napping® combined with verbalization. In terms of statistical analysis, HMFA can perfectly deal with a mix of "pure" Napping® data, "pure" sorting data, and sorted Napping data. Ultimately, we would recommend that subjects use the task they feel most comfortable with, followed by a verbalization phase.

Napping® and sorted Napping as a profiling technique - some complementary thoughts

This section, which is structured into three parts, provides the opportunity to add some additional commentaries on what has already been written in this chapter. The first part contains an overview of the considerable literature relating to Napping®, in which its ability (i.e. reliability) has been evaluated from multiple angles when comparing to other methods, such as QDA® or free sorting. In order to shed light on this issue, the second part will focus on the intrinsic nature of holistic and analytical data. Finally, the third part will point out some differences between free sorting and Napping® – the two holistic methods have been commonly used in sensory analysis.

An overview of Napping®

Napping® is a holistic method that allows collecting the information about the similarity between stimuli. From the perspective of the protocol, PM is very close to Napping®. That is the reason why some authors, such as Nestrud & Lawless (2010) and Hopfer & Heymann (2013), use the two technical terms interchangeably. However, there are other authors, such as Dehlholm et al. (2012) and Giacalone, Ribeiro, & Frøst (2013), who strictly distinguished these two methods. According to these authors, Napping® possesses a specific protocol as it uses a tablecloth 60 cm × 40 cm with white background for data collection, and uses MFA statistical method for data analysis; while PM possesses different protocols as it uses the tablecloth in A4 size (29.7 × 21 cm), A3 (42 × 29.7 cm), or square (60 × 60 cm) with white or grid background for data collection, and uses either PCA, GPA, MDS, or STATIS statistical method for data analysis.

Up to now, the differences between PM and Napping® have not been discussed in terms of the nature of data collection and analysis. Previous studies on Napping® (and PM) focus mainly on its ability. Particularly, the ability of Napping® is evaluated by comparing the product space provided by Napping® to that provided by QDA®. In this context, QDA® is considered a reference method for descriptive analysis. The ability of Napping® is evaluated as high level, meaning that this method is reliable, in the case that its product space is similar to that provided by QDA®; *vice versa*, the ability of Napping® is evaluated as low level, meaning that this method is not reliable, in the case that its product space is different from that provided by QDA®. As seen in the literature, the ability of Napping® (and PM) has been observed and evaluated at different levels.

Risvik et al. (1994) compared three methods PM, QDA®, and (dis)-similarity scaling (DS) on five chocolate products. A panel of 9 trained subjects conducted QDA® using unstructured scale. This panel was then also asked to conduct DS, and PM (a tablecloth in A4 size, white color background marked with two crossed axes). Each method was replicated three times. The data collected by QDA® and PM were analysed by GPA, whereas the data collected by DS were analysed by INDSCAL model. The results showed that the product spaces obtained by the three methods were similar, besides the product spaces obtained solely by the three replications of PM were similar from one to another.

Three years later, Risvik et al. (1997) compared QDA® and PM on 7 blueberry soups. A trained panel of 12 subjects conducted a QDA® (12 sensory attributes with 2 replications). For PM, a panel of 8 consumers was conducted using an A3 white sheet with 3 repetitions. At the end of the PM task, the consumers were asked to record their preference. The data obtained from QDA®, PM, and the preference test were analysed using PCA. The results showed that the consumer panel perceived the samples somewhat differently from the trained panel due to the fact that *"...the best similarity was found when comparing the first dimension, thus suggesting good agreement on the obvious aspects of the product (p. 63)."*

Louw et al. (2013) compared Napping® and Partial Napping to QDA® on 9 brandies. The products were prepared in two sets. The first set consists of 5 commercial products, one product was duplicated. The second set consists of nine commercial products, also one product was duplicated. For tasting sessions, all the samples were diluted to 20% a/v with odourless distilled water. The first panel performed 3 sessions of Napping® and Partial Napping on each of sample set. Three months later, this panel performed QDA® for all the samples. The second panel performed Napping® and Partial Napping on two sample sets without replication. The data obtained from QDA® were analysed by using PCA, and the data obtained by Napping® and Partial Napping were analysed by using MFA. The results showed that Napping® was reliable (its results were similar to those provided by QDA®) and repeatable in a small sample set (about 6 products); whereas Partial Napping could be more stable in a larger sample set (about 10 products).

The "slightly" or "closely" similarity between the results provided by Napping® (or PM) and QDA® leads to the hesitation in applying Napping® for descriptive analysis. This judgment is not completely accurate. In fact, the results provided by Napping® do not need to be similar to those provided by QDA® because Napping® (or PM) collects the holistic data whereas QDA® collects the analytical data. The next section will highlight the differences between holistic and analytical data.

Understanding holistic versus analytical data

A sensory map represents the relative positioning of a set of products from a perceptual point of view. It has numerous applications in various fields, such as: food product development, quality control, understanding consumer preferences, etc. In the field of sensory science, a sensory map can be provided by various sensory profiling techniques, either conventional or rapid profiling techniques. Nevertheless, it is worth noting the two following points:

- Some sensory profiling techniques approach products based on analytical features, while other techniques approach products based on holistic features.
- As the intrinsic nature of the information is totally different from one method to the other, the two methods might lead to different sensory maps: actually, it is expected and in itself is not a problem.

In psychology, analytical and holistic approaches are the two main ways used by the human mind to assess complex objects. The analytical approach considers each part of the object, and the representation of the object is the composition of all parts; whereas, the holistic approach directly grasps the whole without consideration of the parts. The two approaches complement each other; neither approach is good or bad, right or wrong.

In the same manner, in sensory science, when assessing the perception of a set of products, the sensory profiling methods based on analytical approach ask the subject to break down a product into a list of descriptors, then he/she estimates the intensity or existence of each descriptor per product. For the sensory profiling methods based on holistic approach, when assessing the perception of a set of products, a subject considers all descriptors as a whole, and then he/she estimates the similarity between products in terms of the most important descriptors, not all the descriptors.

Let's take a simple example, although somehow "artificial", for illustrating the fact that the two different ways to approach products might lead to two different sensory maps. This example is part of a study, which aims to access consumer-perceived food quality (cf. Appendix 3). In this example, nine stimuli – images of banana – were created according to an experimental design (cf. Table A3 - 1) based on 4 ordered factors with three modalities each.

In terms of analytical data, the table associated with the design could be considered as the sensory profiles of the products: as the factors are ordered, the values in the table could be considered as intensity values (cf. Table 1.4). In terms of holistic data, the stimuli were used in a sorting experiment. The experiment was conducted with 30 subjects.

As usual, analytical data were analysed using Principal Component Analysis (PCA), and as proposed by Cadoret et al. sorting data were analysed using Multiple Correspondence Analysis (MCA).

Table 1.4 Sensory profiles of the products

Product	Colour	Shape	Bruise	Cigar-end rot
1	1	1	1	1
2	1	2	3	2
3	1	3	2	3
4	2	1	3	3
5	2	2	2	1
6	2	3	1	2
7	3	1	2	2
8	3	2	1	3
9	3	3	3	1

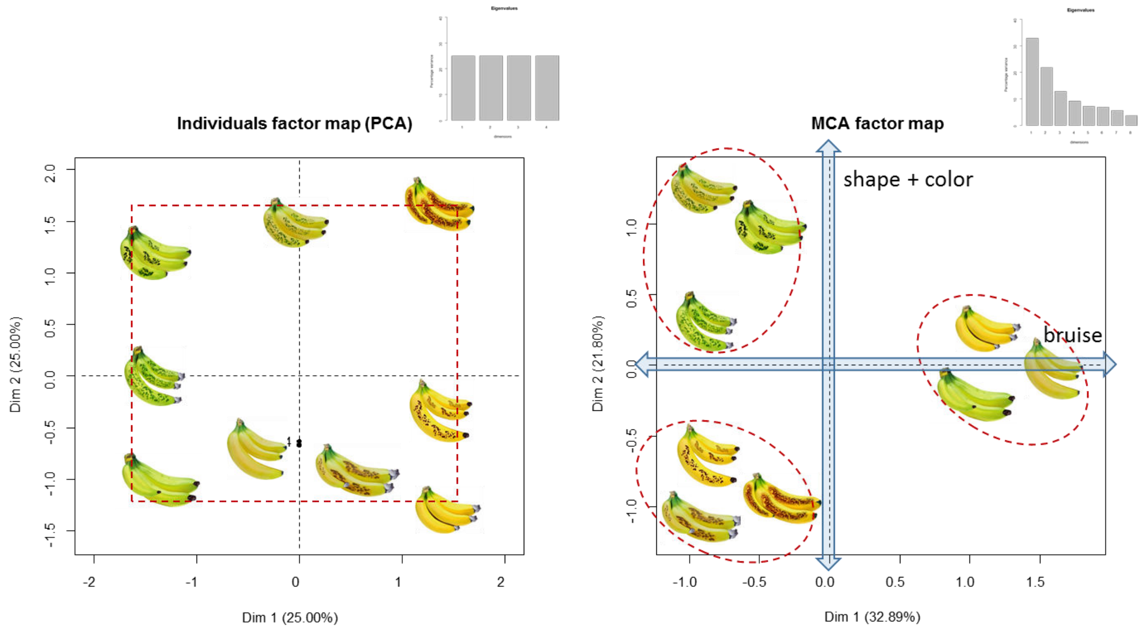


Figure 1.10 Product space obtained by: (a) an experimental design and (b) a sorting task

Figure 1.10 (a) & (b) illustrates the sensory maps obtained from the analytical and holistic approach, respectively. In Figure 1.10.a, the products are distributed dispersedly on the sensory map. To each dimension structuring the sensory map corresponds one of the 4 factors based on which the stimuli were built (Dim.1 – cigar, Dim.2 – color, Dim.3 – shape, and Dim.4 – bruise), each one explaining the same amount of variance (i.e., 25% of the total variance). In other words, descriptors are orthogonal, none is more important than the other.

In Figure 1.10.b, the products are sorted in three groups on the sensory map. Dim.1 opposes the "bruise" bananas to the "non-bruise" bananas, and Dim.2 opposes "green-colour and homogeneity-shape" bananas to "yellow-colour and high heterogeneity-shape" bananas. These results showed that subjects perceived the bruise factor as an important factor of variability among the bananas, which explains 32.89% of the total variance.

The sensory map obtained by holistic approaches is not more interesting than the one obtained by analytical approach (or vice versa). Analytical and holistic approaches are two different ways to assess products, they complement each other, and the choice of the approach depends on the purpose of the studies: whether an accurate sensory description is needed, or a sensory description based on a consumer point of view.

Understanding Napping® versus sorting

For purposes of illustration, throughout this paragraph, sorting and Napping® will be assimilated to cameras which purpose is to provide a picture of a set of stimuli. For this analogy, a subject will act as a photographer.

The common point between these two methods is the fact that subjects assess a set of stimuli through their own prism, using their own camera angle, on the basis of their perceived similarities (hence the concept of similarity-based methods; Valentin, Chollet, Lelièvre, & Abdi, 2012). In other words, the picture taken – the partition on the stimuli in the case of sorting, the positioning of the stimuli in the case of Napping® – depends on the point of view chosen by the photographer (by definition, the spot where the picture is taken) and the focal length of the optical system of the camera.

Sorting is certainly the most frequently used holistic method. One of the reasons could be that the sorting procedure is simple, as it is based on categorisation – a natural cognitive process routinely used in everyday life (Chollet et al., 2011). Subjects are asked to sort products in groups in the condition that: products are assigned in the same group if they are perceived as similar, and in different groups if they are perceived as different. Very often, the number of groups is not limited, and generally to each group corresponds a personal criterion perceived by one subject. This particular feature of the task impacts directly the optical system of the “sorting” camera: technically, the camera is equipped with a zoom lens that enables the photographer to change his/her focal distance. The picture can be taken using a small focal distance and in that case the angle will be large, the number of groups will tend to be small. On the contrary, the picture can be taken using a long focal distance and in that case the angle will be narrow, the number of groups will tend to be large.

In Napping®, the 60 by 40 cm rectangular on which the stimuli are positioned impacts also the optical system of the “Napping®” camera: technically, the camera is equipped with a prime lens which is a length whose focal length is fixed. Somehow, the two dimensions of the rectangle are imposing a focal distance to the photographer, in other words its forces the photographer to have some perspective on the whole set of stimuli.

To conclude, being equipped with a zoom lens or a prime lens is neither good nor bad. One could advocate that the more degrees of freedom, the better, with the risk that photographers use different distances and therefore with the risk that the common picture might be unstable. One could advocate that a fixed focal distance is too constraining, with the advantage that photographers use the same angle.

(This page is left blank)

2

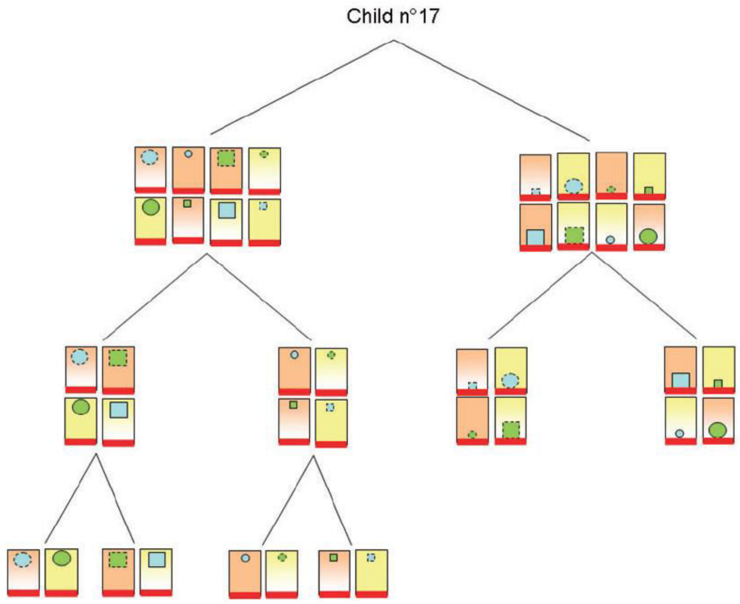
A Napping® based methodology to quickly obtain a model of emotions

A first attempt to integrate subject's cognitive process based on a model approach

Ce chapitre 2 constitue une suite logique du chapitre 1 qui nous a sensibilisés à l'importance de prendre en compte le comportement du sujet lors d'une tâche de Napping® : cette prise en compte a abouti à une variante du Napping® appelée Napping catégorisé.

Dans le premier chapitre, alors que l'information intégrée découle de la configuration finale des stimuli fournie par un sujet, il nous a semblé intéressant ici de réfléchir à l'intégration d'une information intermédiaire, entre le moment où le sujet commence la tâche et le moment où il la termine. Cette information serait constituée des étapes clés par lesquelles passe un sujet pour aboutir à sa configuration finale. Ce type d'information, qui peut être assimilé au processus cognitif du sujet, a déjà été abordé par Cadoret et al. (2011) dans le cadre du tri hiérarchique. Dans ce cadre, les auteurs proposent d'utiliser la suite de partitions emboîtées issue du tri hiérarchique d'un sujet pour représenter le processus cognitif de ce même sujet. Dans leur article, ces auteurs se posent la question de la multiplicité des processus cognitifs lors d'une tâche de tri hiérarchique pour aboutir à une même partition sur les stimuli (cf. Figure 2.1).

L'expérience mise en œuvre pour illustrer leur méthodologie consistait à proposer une tâche de tri hiérarchique à des enfants de 7 à 10 ans. Il s'agissait, pour 89 enfants, de trier un ensemble de 16 cartes construites sur la base d'un plan d'expérience à 7 facteurs à deux modalités. L'objectif de l'expérience était d'identifier les traits saillants perçus par les enfants. Il a été constaté, grâce à la représentation des processus cognitifs, que des enfants pouvaient obtenir une même partition finale sur les stimuli, sans pour autant passer par les mêmes étapes de tri, autrement dit, sans pour autant raisonner de la même façon.

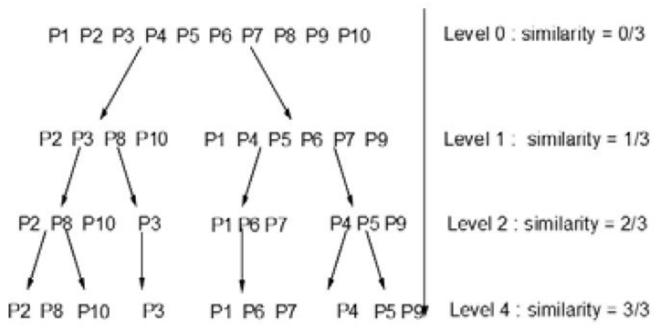


	L1	L2	L3
A	Group1	Group1	Group5
B	Group2	Group3	Group3
C	Group1	Group2	Group6
D	Group1	Group1	Group1
E	Group1	Group2	Group6
F	Group1	Group1	Group5
G	Group2	Group3	Group3
H	Group2	Group4	Group4
I	Group2	Group3	Group3
J	Group1	Group2	Group2
K	Group2	Group3	Group3
L	Group2	Group4	Group4
M	Group1	Group1	Group1
N	Group1	Group2	Group2
O	Group2	Group4	Group4
P	Group2	Group4	Group4

Figure 2.1 New way for coding descending hierarchical sorting data (Cadoret et al., 2011, p. 98)

Curieusement, malgré le potentiel d'utilisation de cette représentation, l'étude du processus cognitif d'un sujet lors d'une tâche de tri hiérarchique n'a pas été abordée par d'autres chercheurs : Valentin et al. (2012) examinent le traitement de données issues d'un tri hiérarchique à travers la constitution d'une matrice de distance entre stimuli (cf. Figure 2.2); et Santosa, Abdi, & Guinard (2010) étudient le lien entre la représentation des sujets et celle des stimuli, mais ne s'intéressent pas au processus cognitif en tant que tel (cf. Figure 2.3).

L'objectif de ce chapitre est de proposer un modèle du processus cognitif de chaque sujet, basé sur sa configuration finale. Ce modèle repose sur une hypothèse forte, discutable, mais néanmoins raisonnable : face à un grand nombre de stimuli, un sujet adopte une approche descendante (top-down) qui consiste à décomposer un ensemble en éléments toujours plus détaillés. Cette hypothèse adoptée, on adaptera la démarche proposée par Cadoret et al. à la tâche du Napping®.



	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
P1	1	0	0	0.33	0.33	1	1	0	0.33	0
P2	0	1	0.33	0	0	0	0	1	0	0.66
P3	0	0.33	1	0	0	0	0	0.33	0	0.33
P4	0.33	0	0	1	0.66	0.33	0.33	0	0.66	0
P5	0.33	0	0	0.66	1	0.33	0.33	0	1	0
P6	1	0	0	0.33	0.33	1	1	0	0.33	0
P7	1	0	0	0.33	0.33	1	1	0	0.33	0
P8	0	1	0.33	0	0	0	0	1	0	0.66
P9	0.33	0	0	0.66	1	0.33	0.33	0	1	0
P10	0	0.66	0.33	0	0	0	0	0.66	0	1

Figure 2.2 Illustration of data coding in descending hierarchical sorting (Valentin et al., 2012, p. 7)

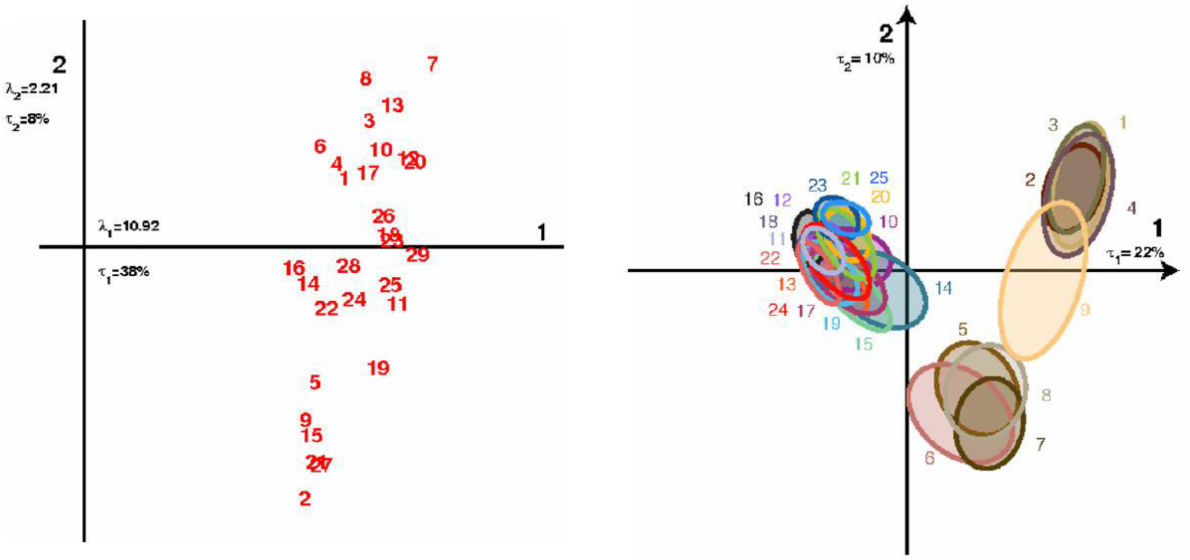


Figure 2.3 Representation of the subjects' map and products' map in descending hierarchical sorting (Santosa, Abdi, & Guinard, 2010, p. 886)

Ce chapitre correspond à l'article "A Napping® based methodology to quickly obtain a model of emotions" soumis à la revue *Behaviour Research Methods*. Dans ce contexte, il nous a semblé important d'équilibrer tant la méthode qui aborde la notion de modèle d'un processus cognitif lors d'une tâche de Napping®, que les résultats qui portent sur un modèle sur les émotions. Dans une première lecture, indépendamment de l'intérêt qu'il peut porter aux émotions, on demandera au lecteur de se focaliser sur la notion de processus cognitif, que l'on retrouvera tout le long de la thèse et que l'on discutera à la fin de l'article notamment.

Chapter outline

This chapter includes the publication:

Lê, M.T., & Lê, S. (2014, submitted). A Napping® based methodology to quickly obtain a model of emotions, Behaviour Research Methods.

Abstract

In this paper, we propose a methodology to obtain an empirical model of emotions[1]. This methodology comprises the way data are collected and analysed. In terms of results, the model integrates two levels of analysis. The first one, based on a consensus over the subjects that have been questioned, provides a multidimensional representation, and a typology of the emotions. This level appears to be of interest to psychologists, whether they adopt the dimensional or the categorical approach. The second one, which represents each subject in relation to the way he constructs his own emotional space. This level seems to be of interest to those cognitivists who are interested in the cognitive process and how it can lead to a structure on emotions.

Napping® combined with Multiple Factor Analysis and Hierarchical Divisive Analysis constitutes the core of this methodology, which is illustrated through a study carried out on 121 Vietnamese subjects and a selection of 53 emotions.

Keywords: Empirical model of emotions; Subjective cognitive process; Napping®.

[1] In the context of this chapter, a model of emotions can be defined as the way emotions are structured.

2.1 Introduction

Although the relationships between emotions have been heavily studied for decades, this research topic remains very much alive today. As a matter of fact, as mankind keeps on evolving, the way emotions are experienced and perceived may probably evolve as well. To conceptualize the structure among emotions or emotional states, many points of view have been proposed so far, amongst which we can quote the two most important ones in terms of influence, the categorical approach and the dimensional approach.

According to the categorical approach, emotional processes can be explained by a small set of basic or fundamental emotions that are innate, not reducible and common to all mankind. Those emotions would be perfectly discerned in memory, in a categorical way, and would be associated with information, ideas and memories also organized in a categorical manner. This approach was developed from a biological perspective by incorporating an evolutionary standpoint from the work of Darwin (1872/1998). Emotions are seen as a means to regulate behavior.

The dimensional approach has been developed in parallel with the categorical approach, often in order to compensate for the shortcomings of the latter. According to the dimensional approach, emotions can be decomposed into two, three, or sometimes four primary factors. The aim of this approach is to draw an analogy between the way emotions are structured and the way they are experienced.

From these points of view, researchers suggest models that allow apprehending relations between emotions. Those models are either issued from pure theoretical considerations (e.g., Russell, 1980; Watson & Tellegen, 1985; Larsen & Diener, 1992; Barrett, 1998, for the dimensional approach; and Ekman & Cordaro, 2011; Izard, 2011; Levenson, 2011; Panksepp & Watt, 2011, for the categorical approach) or from results based on empirical data (e.g., Fontaine, Scherer, Roesch, & Ellsworth, 2007, for the dimensional approach; and Shaver, Schwartz, Kirson, & O'Connor, 1987, for the categorical approach). Tracy & Randles (2011) compare four theoretical models issued from the categorical approach. Nevertheless, they "leave to future researchers the difficult empirical work of determining which pieces of each model are most correct" (Tracy & Randles, 2011, p. 397). It is partly in this context that our research may be positioned.

The aim of our paper is to propose a new methodology for obtaining an empirical model of emotions, based on the way they are perceived by subjects. In terms of results, the model integrates two levels of analysis.

The first one, based on a consensus over the subjects that have been questioned, provides a multidimensional representation, and a typology of emotions. This level appears to be of interest to psychologists, whether they adopt the dimensional or the categorical approach.

The second one represents each subject in relation to the way he constructs his own emotional space. This level seems to be of interest to those cognitivists who are interested in the cognitive process and how it can lead to a structure on emotions.

The Napping® method for the data collection combined with Multiple Factor Analysis and Hierarchical Divisive Analysis for the data analysis constitutes the core of this methodology. This methodology is illustrated through an empirical study carried out on 121 Vietnamese subjects and a selection of 53 emotions.

2.2 Methodology

2.2.1 Napping®

Projective methods, also called projective techniques, encompass a large family of personality tests. Such techniques allow people to reveal their personalities by experiencing different stimuli such as emotions (Donoghue, 2000; Steinman, 2009).

From an experimental point of view, two reference methods are mainly used to collect data in order to obtain a representation of the way emotions might be structured. The first one consists of asking subjects to gather stimuli into clusters according to the way they perceive their similarities – free sorting task (Russell, 1980; Shaver et al., 1987; Storm & Storm, 1987; Thomson & Crocker, 2013). The second one consists of asking subjects to provide a distance between two stimuli according to the way they perceive their similarities – pairwise similarity rating (Niedenthal & Halberstadt, 2003; Tsuji, Shimokawa, & Okada, 2010).

The first method may not be fully satisfactory because relationships can exist between clusters of stimuli (e.g., an order relation): this information is not accessible due to the fact that it is never asked of subjects. The second method may show itself to be unfeasible due to the number of pairs of stimuli to assess.

Pagès (2005) proposed a methodology called Napping® (in French, the word "nappe" means sheet or tablecloth). This projective technique consists in asking subjects to position stimuli on a 60 cm by 40 cm rectangle, according to the way they perceive their similarities. Practically, this method shows itself to be very advantageous.

Firstly, without being asked, subjects are prioritizing the reasons why they perceive the stimuli differently: they are naturally influenced by the X-axis, which is by construction 1.5 more important than the Y-axis of the rectangle. In other words, the rectangle in which stimuli are positioned forces subjects to privilege one dimension first (i.e., the X-axis) and to separate stimuli with respect to that dimension, then, to use the second dimension (i.e., the Y-axis). Secondly, if people decide to cluster stimuli, we may easily exploit the information concerning the relative position of one cluster with respect to another cluster. Finally, it is less fatiguing for measuring absolute differences between two stimuli, as the set of stimuli is perceived as a whole from a global point of view, contrary to pairwise comparisons from a more local point of view.

2.2.2 Statistical analysis

2.2.2.1 Structuring emotions in dominant dimensions.

As a result of the Napping® experience, each subject provides two vectors of coordinates of dimension $I \times 1$ each (one for the X-axis, one for the Y-axis), where I denotes the number of stimuli to be positioned on the rectangle (cf. Figure 2.4.a & b). Hence, the final data set to be analysed, denoted X, is obtained by merging the N couples of vectors of coordinates, where N denotes the number of subjects (cf. Figure 2.4.c). In other words, *X* can be seen as a data set structured into *N* groups of two variables each. Typically, the statistical analysis of such data set *X* should take into account the "natural" partition on the variables.

Multiple Factor Analysis (MFA) (Escofier & Pagès, 1994; Pagès, 2015) was conceived precisely for the purpose that the part of each group of variables be balanced within a global analysis. When variables are quantitative, MFA can be defined as an extension of Principal Component Analysis (PCA) that takes into account a group structure on the variables and that balances the part of each group: in that sense, MFA can be seen as a weighted PCA. Technically, each group of variables is weighted by the inverse of its highest eigenvalue.

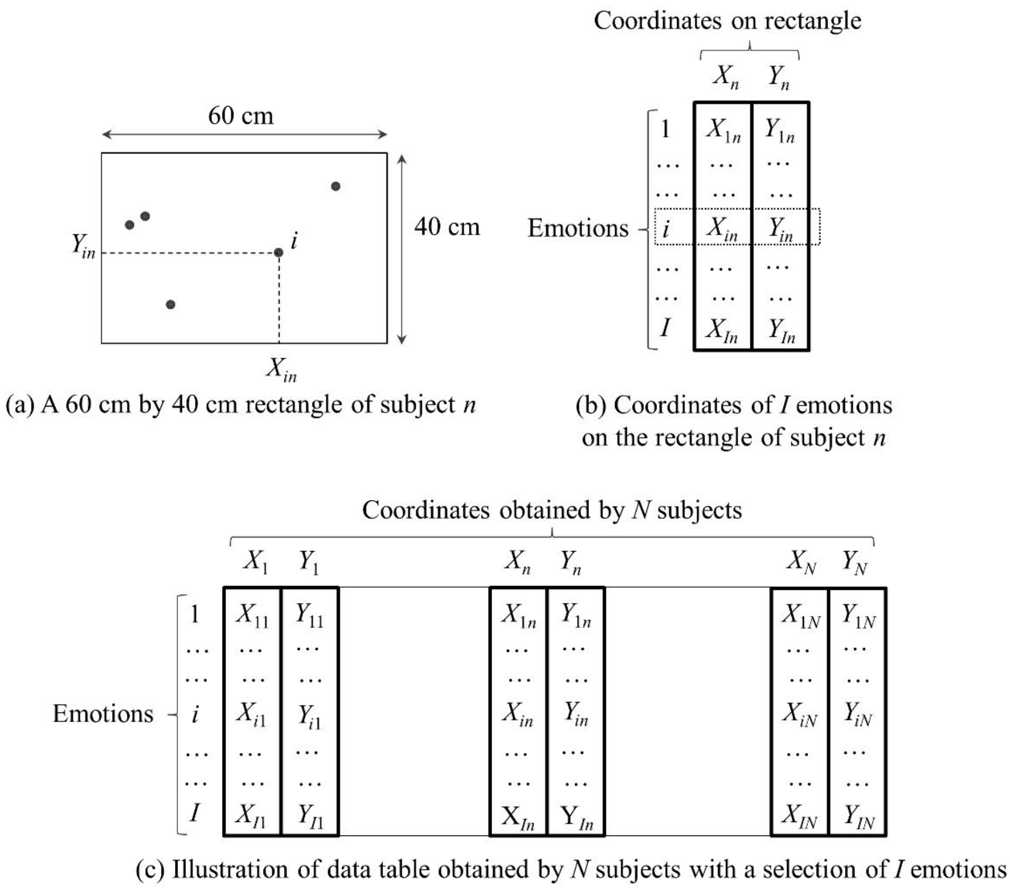


Figure 2.4 Illustration of data table obtained by the Napping® method

The rationale behind this weighting scheme is to exhibit dimensions that are (1) common to as many groups of variables as possible, (2) specific to some groups of variables. From this perspective, MFA can be defined as a variant of generalized canonical analysis (Carroll, 1968).

Although researchers have almost exclusively focused on the first two dimensions of the emotional space, the research question concerning the number of significant dimensions remains open. Complementary to the interpretation, we use permutation based methods to assess the number of significant dimensions (cf. Appendix A2.1).

2.2.2.2 Structuring emotions in clusters

Unsupervised classification is actually complementary to exploratory multivariate methods (Husson et al., 2011; Lebart, Morineau, & Warwick, 1984): once the dominant dimensions are identified, it is often interesting to revisit the information at a cluster level. In order to fit in with the categorical approach, we chose, amongst the different cluster analysis methods, hierarchical clustering (versus direct partitioning). Indeed, similar to the categorical approach, which structures emotions as a hierarchy of clusters of emotions, hierarchical clustering seeks to build a hierarchy of clusters.

There are mainly two types of hierarchical clustering algorithms: the ones that are qualified as agglomerative in the sense that clusters are merged iteratively (from the singletons to the whole set); the ones that are qualified as divisive in the sense that clusters are divided into two sub-clusters iteratively (from the whole set to the singletons). We will see in the next paragraph that this choice is determinant for studying the cognitive process of the subjects.

In the literature, hierarchies can be obtained using either a bottom-up strategy as in Thomson & Crocker (2013), or a top-down strategy as in Laros & Steenkamp (2005). Our choice of using a divisive algorithm is mainly based on the fact that we want to mimic subjects when they are confronted with a Napping® task. Detailed observations were carried out on subjects performing a Napping® task and revealed that subjects were almost exclusively using a top-down strategy. As evoked previously, because stimuli are positioned within a rectangle, they are first separated with respect to the length, then with respect to the height, in a hierarchical way.

Divisive ANAlysis (DIANA), which is based on a divisive algorithm proposed by MacNaughton-Smith, Williams, Dale, & Mockett (1964), is used to obtain a hierarchy from the whole set of emotions to the singletons. This algorithm is applied on the results issued from MFA, in other words on the representation of the emotions based on all the subjects.

2.2.2.3 Representing a cognitive process as a hierarchy.

Let us first define, very restrictively, what we call cognitive process within our context. With respect to the way data are collected, the cognitive process of a given subject can be defined as the sequence of steps that has led the subject to his stimuli configuration on the rectangle (in other words, to his couple of vectors of coordinates).

Similarly to what has been explained in the previous paragraph, to obtain this sequence of steps, we propose to apply the DIANA algorithm on each subject configuration. Indeed, the main result of this algorithm is a hierarchy, and a sequence of steps, as a sequence of nested partitions, can be assimilated to a hierarchy. In Figure 2.7.a, for instance, the cognitive process associated with subject 1 is represented by a hierarchy that first separates the emotions into two clusters (i.e., the first level of the hierarchy) the positive emotions and the negative ones. Then, the second step of his cognitive process (i.e., the second level of the hierarchy) separates the negative emotions into two clusters of emotions, described as activated and deactivated, and so on. Each step of the cognitive process can be assimilated to a partition on the stimuli. As such, it can be assimilated to a categorical variable: in Figure 2.7, for instance, the first step can be considered as a categorical variable with two categories, a first one that gathers the positive emotions, and a second one that gathers the negative ones.

Thus, a hierarchy, as a sequence of steps, is represented through its sequence of categorical variables. Hierarchies are represented within the representation of the stimuli based on all subjects, obtained thanks to MFA as explained in the section 2.2.2.1. To do so, we represent each hierarchy through its sequence of categorical variables, on the axes issued from MFA: the coordinate of a categorical variable j on the axis F_s of rank s is calculated using the correlation ratio (Escofier & Pagès, 2008, p.98) (cf. Appendix A2.2). As all cognitive processes are represented within the same referential, they can be easily compared.

2.2.3 Presentation of the case study

To illustrate our methodology, we present results issued from a case study. For this experiment, 121 Vietnamese subjects were asked to position 53 emotions (cf. Table 2.1) on a rectangle where the ratio of its length by its width equals 1.5. Most of the subjects were students at Ho Chi Minh City University of Technology, and were 18 to 25 years old.

Table 2.1 List of emotion terms used in this case study.

Vietnamese	English	Vietnamese	English
giàu tình cảm	affectionate	ghen tỵ	jealous
giận dữ	angry	vui sướng	joyful
tức tối	annoyed	lười biếng	lazy
xấu hổ	ashamed	cô đơn	lonely
yêu mến	beloved	mất phương hướng	lost
chán nản	bored	luyến tiếc	nostalgic
điềm tĩnh	calm	lạc quan	optimistic
bối rối	confused	hoảng sợ	panicked
dũng cảm	courageous	say mê	passionate
hiếu kỳ	curious	yên ổn	peaceful
thất vọng	disappointed	đáng tiếc	pitiful
nghi ngờ	doubtful	hài lòng	pleased
háo hức	eager	hãnh diện	proud
ngượng ngùng	embarrassed	ân hận	regretful
điên dại	enraged	từ chối	rejected
nhiệt tình	enthusiastic	nhẹ nhõm	relieved
công bằng	fair	buồn	sad
sợ hãi	fearful	thỏa mãn	satisfied
tự do	free	an toàn	secure
thân thiện	friendly	ích kỷ	selfish
phẫn nộ	furious	nhút nhát	shy
cảm giác tội lỗi	guilty	kinh ngạc	stunned
hạnh phúc	happy	ngạc nhiên	surprised
đầy hi vọng	hopeful	tin cậy	trustworthy
tủi nhục	humiliated	lo lắng	worried
bị tổn thương	hurt	trẻ trung	youthful
thờ ơ	indifferent		

The emotions were chosen within a product development context for a consumer study, in order to obtain a list that would be as comprehensive as possible. Pre-tests were conducted to make sure that the list was easily understandable. Beyond the methodology presented in this article, the main objectives of this case study were (1) to understand how emotions were perceived and structured, (2) to find a way to extract a subset of emotions.

Practically, data were collected on a computer screen, via a VBA application developed by the Applied Mathematics Department of Agrocampus Ouest, France. The application is freely available at the following address: http://napping.agrocampus-ouest.fr.

Data were analysed using the R software version 3.1.1 (R Core Team, 2014). Multiple Factor Analysis and the DIANA algorithm for Hierarchical Divisive Clustering were carried out with the FactoMineR package version 1.27 (Husson, Josse, Lê, & Mazet, 2014; Lê, Josse, & Husson, 2008) and the cluster package version 1.15.2 (Maechler, Rousseeuw, Struyf, Hubert, & Hornik, 2014), respectively.

2.3 Results

The main objective of this paragraph is to illustrate how the outputs provided by our methodology can be exploited by the end-user. Therefore, we provide general rules for interpreting the results mentioned in section 2.2.2. A finer interpretation of the results, in a given context, should be left to the specialists.

2.3.1 Structuring emotions in dominant dimensions

Figure 2.5 corresponds to the representation of the individuals (i.e., emotions) obtained by using MFA. Here, MFA is performed on groups of two variables each (i.e., the X and Y coordinates of the stimuli on each subject's rectangle).

This representation is the consensus of the subjects on the first factorial plane, constituted by the first two main dimensions of variability. The first dimension, denoted Dim 1 in Figure 2.5, explains 29.52% of the total variance. This relatively high value (considering the number of individuals and variables in the analysis), indicates that there is an important consensus among subjects.

This first dimension clearly opposes emotions such as “happy” and “joyful” (on the left side of the plane), that could be qualified as positive, to emotions such as “lost” and “sad” (on the right side of the plane), that could be qualified as negative.

The second dimension, denoted Dim 2 in Figure 2.5, explains 6.38% of the total variance. Although this value seems apparently low, this dimension will be kept in the analysis since it has been tested as statistically significant (cf. Appendix A2.3), and obviously interesting in terms of interpretation.

Amongst the so-called negative emotions, Dim 2 opposes emotions such as “angry” and “furious”, that could be qualified as activated, to emotions such as “guilty” and “regretful”, that could be qualified as deactivated.

Remark 1: Let us notice the singular position of the two emotions “stunned” and “surprised” located near the centre of gravity of the plane. It seems that they have been isolated by the subjects, and that they have been perceived as neither positive nor negative: one can be “stunned” or “surprised” in either positive or negative ways; “stunned and “surprised” could easily be qualified as “neutral”.

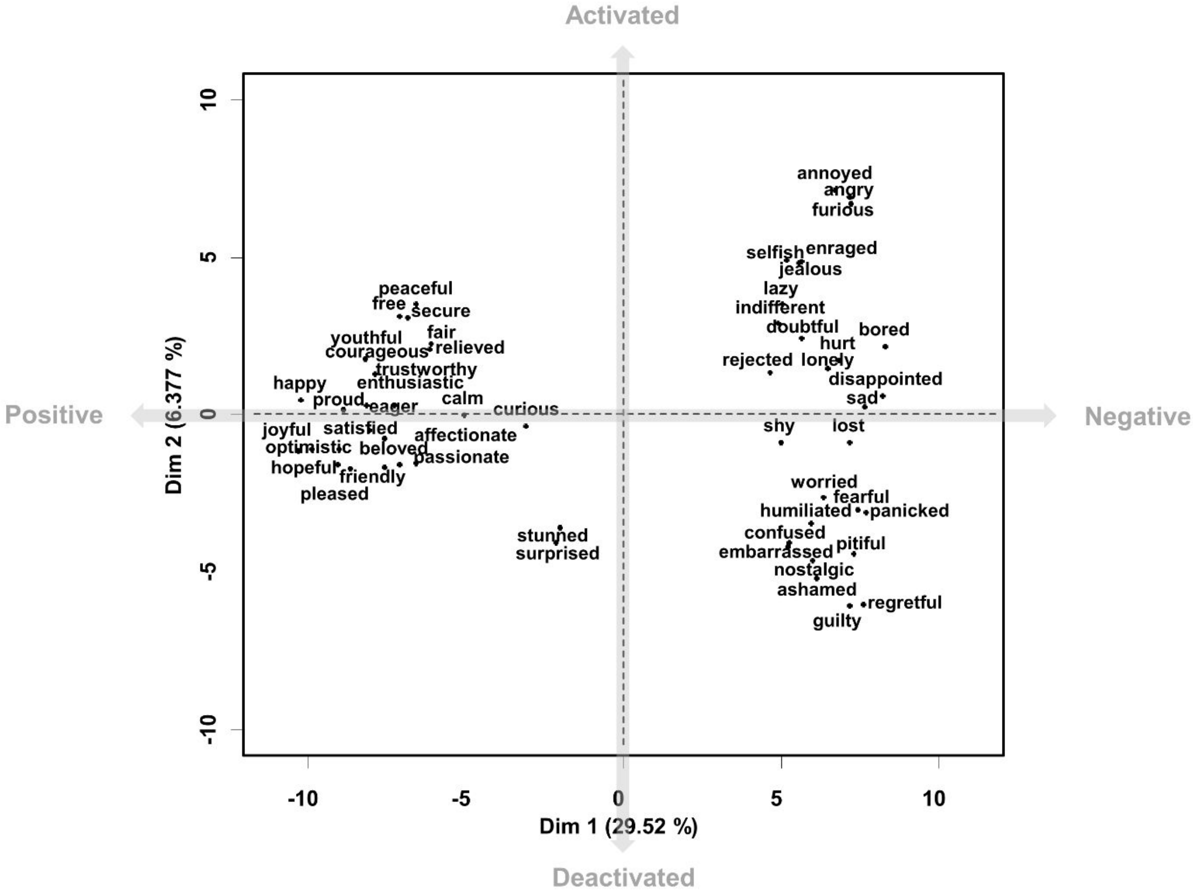


Figure 2.5 Structure of emotions in the first factorial plane obtained by MFA.

Remark 2: In terms of dispersion, we can see the difference between the cluster of so-called positive emotions and the cluster of so-called negative emotions: positive emotions seem to be more homogeneous than negative emotions, which are more heterogeneous. This may lead to many different interpretations that should be left to the specialists. One of them, could be that, with respect to the subjects, positive emotions (when merged with negative emotions) are first perceived as a whole, which may be due to their strong common character (i.e., positiveness) perceived by the subjects; whereas negative emotions (once isolated from the positive emotions) are perceived more like some kind of continuum. Such evidence has already been observed in Shaver et al. (1987, Figure 2, p. 1069), Fontaine et al. (2007, Figure 1.a, p. 1055) and Chrea et al. (2009, Figure 1, p. 54).

Remark 3: The solution provided by MFA is not only 2-dimensional. As evoked later in the discussion, the analysis of the third dimension provides meaningful information related to positive emotions, despite its seemingly low percentage of variability, low yet still significant (cf. A2.3).

2.3.2 Structuring emotions in clusters

Figure 2.6 corresponds to the representation of the individuals (i.e., emotions) obtained by using the divisive clustering algorithm DIANA. Here, DIANA is performed on the distance matrix between emotions resulting from the previous MFA, in which the part of each subject's rectangle has been balanced.

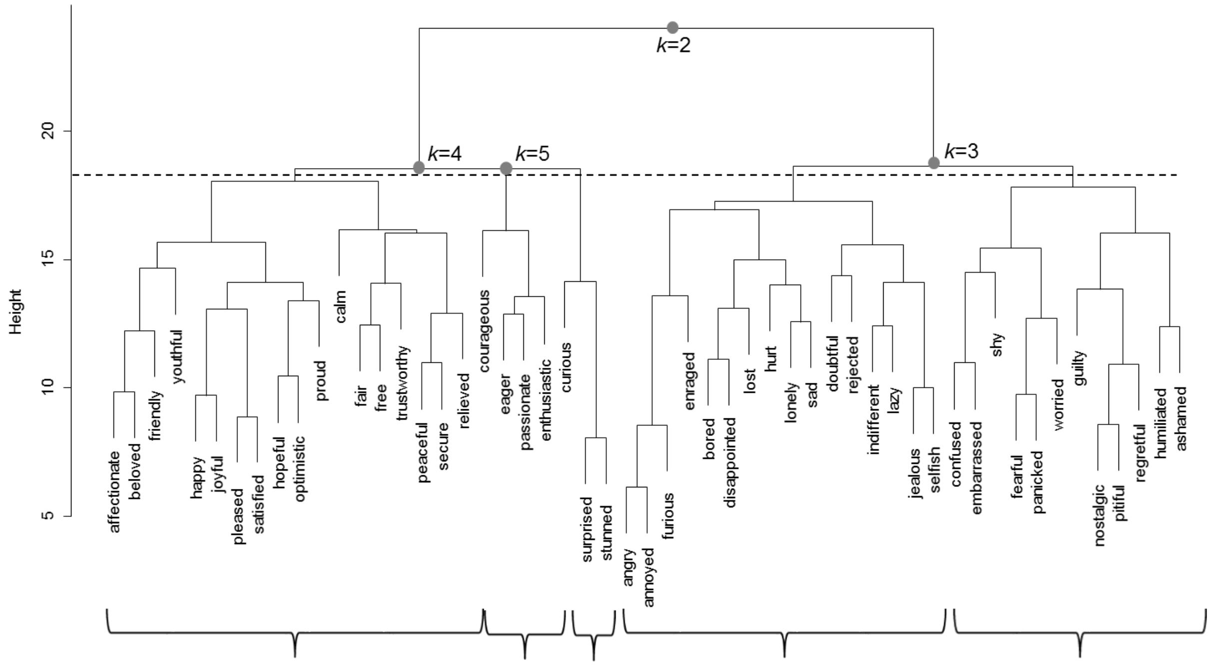


Figure 2.6 Hierarchy at the panel level obtained by divisive algorithm.

This representation depicts a hierarchy of clusters of emotions, in other words a sequence of nested partitions on the emotions. Starting from the top of the hierarchy to the bottom, the number of clusters per partition increases one by one. The resulting hierarchy can be mainly traversed in two ways: (1) by considering each level separately, which enables us to obtain a typology of emotion as presented below; (2) by considering the levels sequentially, which enables us to assess the cognitive process of subjects as presented in the next section.

As it is usually the case for clustering methods, the end-user has to determine a cutting point of the hierarchy which will lead to a typology associated with a certain number of clusters. For example, as illustrated by the broken line in Figure 2.6, the hierarchical level when the number of clusters, denoted k, equals 5. On this node, the emotions within each cluster can be determined and summarized as in Table 2.2.

Remark 4: The typology of emotions, when k=5, seems to be mostly coherent with the dispersion among emotions in the first factorial plane. In this, cluster 1 and cluster 2 consist of all negative emotions. With the continuum of decreasing intensity, cluster 1 (e.g., "angry", "annoyed") is qualified as negative-activated cluster; whereas cluster 2 (e.g., "regretful", "guilty") is qualified as negative-deactivated cluster. Secondly, cluster 3 consists of "curious", "stunned", and "surprise" which are positioned in the centre of the factorial plane, and now are grouped together in one cluster. As mentioned before, it seems that these emotions have been perceived as neither positive nor negative, they are qualified as a neutral cluster. And finally, cluster 4 and cluster 5 consist of all positive emotions which are homogenous in the first factorial plane, but divided into two separate clusters in this section by the algorithm.

The dispersion of cluster 4 and 5 will be taken into account in other factorial planes. Furthermore, the interpretation of the typology of five clusters, which is in accordance with Vietnamese culture, will be presented in the Discussion section.

Table 2.2 The typology of five clusters of emotions obtained by divisive algorithm.

Cluster 1: angry, annoyed, furious, enraged, bored, disappointed, lost, hurt, lonely, sad, doubtful, rejected, indifferent, lazy, jealous, selfish
Cluster 2: confused, embarrassed, shy, fearful, panicked, worried, guilty, nostalgic, pitiful, regretful, humiliated, ashamed
Cluster 3: curious, surprised, stunned
Cluster 4: courageous, eager, passionate, enthusiastic
Cluster 5: affectionate, beloved, youthful, happy, joyful, pleased, satisfied, hopeful, optimistic, proud, calm, fair, free, trustworthy, peaceful, secure, relieved, friendly

2.3.3 Representing a cognitive process as a hierarchy

Figure 2.7 illustrates the hierarchies of subjects obtained by using the divisive algorithm. The divisive algorithm enables us to study the dynamic of the hierarchy, due to the fact that it focuses on the process of successive splitting clusters rather than on the partitions at each hierarchical level.

The interpretation of Figure 2.7 mimics the way the DIANA algorithm works from top to bottom. Here, the algorithm DIANA is performed on the distance matrix among emotions provided by each subject, in other words, on the distance matrix based on each subject's rectangle.

The four subjects represented in this figure were chosen as they correspond to four different ways of positioning emotions on the rectangle (cf. Figure A-3), which have led to four different strategies for clustering emotions. In order to comprehend the cognitive processes of such subjects, we mimic and illustrate their first steps. The number of steps to illustrate our methodology was arbitrarily chosen equal to 4 (k=5).

In Figure 2.7.a, at the first step (when k=2), subject 1 focuses on valence by the fact that emotions are divided into two clusters qualified as positive and negative. Then, at the second step (when k=3), the negative emotions are divided into two clusters qualified as activated and deactivated. At the third and fourth steps (when k=4 and k=5), only the positive emotions are divided.

In Figure 2.7.b, at the first step (when k=2), subject 11 divides emotions into positive and negative clusters. At the second step (when k=3), the positive emotions are divided into two finer clusters, but not the negative emotions as in the hierarchy of subject 1. Then, at the third step (when k=4), the negative emotions are divided. And finally, at the four step (when k=5), only one of the positive clusters is divided.

In Figure 2.7.c, at the first step (when k=2), subject 73 divides emotions into activated and deactivated clusters. One of the two clusters seems to be qualified as an activated cluster due to the fact that it contains most of the negative activated emotions (cf. Cluster 1 in Table 2.2) such as "angry", "furious", "enraged", "annoyed", "selfish", "indifferent", "doubtful", and "jealous". However, there are also some of the positive emotions such as "happy", "relieved", "peaceful", and "secured" which evoke a feeling of well-being and security, hence could be considered to be "activated". Notably, this first step of subject 73 is not the same as the one of subject 1 or subject 11. For the continuation, only the deactivated cluster is divided.

In Figure 2.7.d, for all steps (from k=2 to k=5), the way of positioning emotions on the rectangle of subject 59 is totally different compared to that of the subjects 1, 14 and 73. In other words, subject 59 uses a different strategy for clustering emotions, neither positive-negative nor activated-deactivated.

In the next paragraph, we will present how the cognitive process of subjects can be graphically represented in the emotional space (e.g., the first factorial plane) in order that we can compare the cognitive process among subjects. In our context, the cognitive process corresponds to a sequence of steps that leads a subject to his stimuli configuration on the rectangle. This sequence of steps per subject can be summarized by using a hierarchy as in Figure 2.7.

Figure 2.8 represents the cognitive process of the subjects in Figure 2.7. To obtain this figure, the steps of a given subject are first represented as explained in Appendix A2.2, by calculating the correlation ratio between a step considered as a categorical variable and a vector of coordinates induced by MFA; then the steps are linked by lines in chronological order, from the first step (with two clusters of emotions) to the last one (with five clusters of emotions). By definition, the value of the correlation ratio lies between 0 and 1. When the coordinate of a step for a given subject on a dimension equals 1, it means that the structure induced by the groups provided by this subject is perfectly consistent with the structure induced by the dimension. On the contrary, when the coordinate of a step for a given subject on a dimension equals 0, it means that the structure induced by the clusters provided by this subject has nothing to do with the structure induced by the dimension. Let's recall that dimensions are provided by all subjects (i.e., the whole panel); therefore this graphical representation allows the comparison of each subject with the whole panel in one part, and of each subject with each other in other part.

For instance, at the first step of subject 1 where the number of clusters k equals 2, its coordinate (denoted $k2s1$) takes a rather high value on the first dimension since the two clusters provided by subject 1 at this step induce the same structure as the whole panel on the first dimension by opposing negative emotions to positive ones. As expected, the four different hierarchies led to four different cognitive process representations.

Remark 5: A cognitive process can be represented by a sequence of more than four steps.

Remark 6: A cognitive process can be projected on different factorial planes which can be constituted by all the significant dimensions.

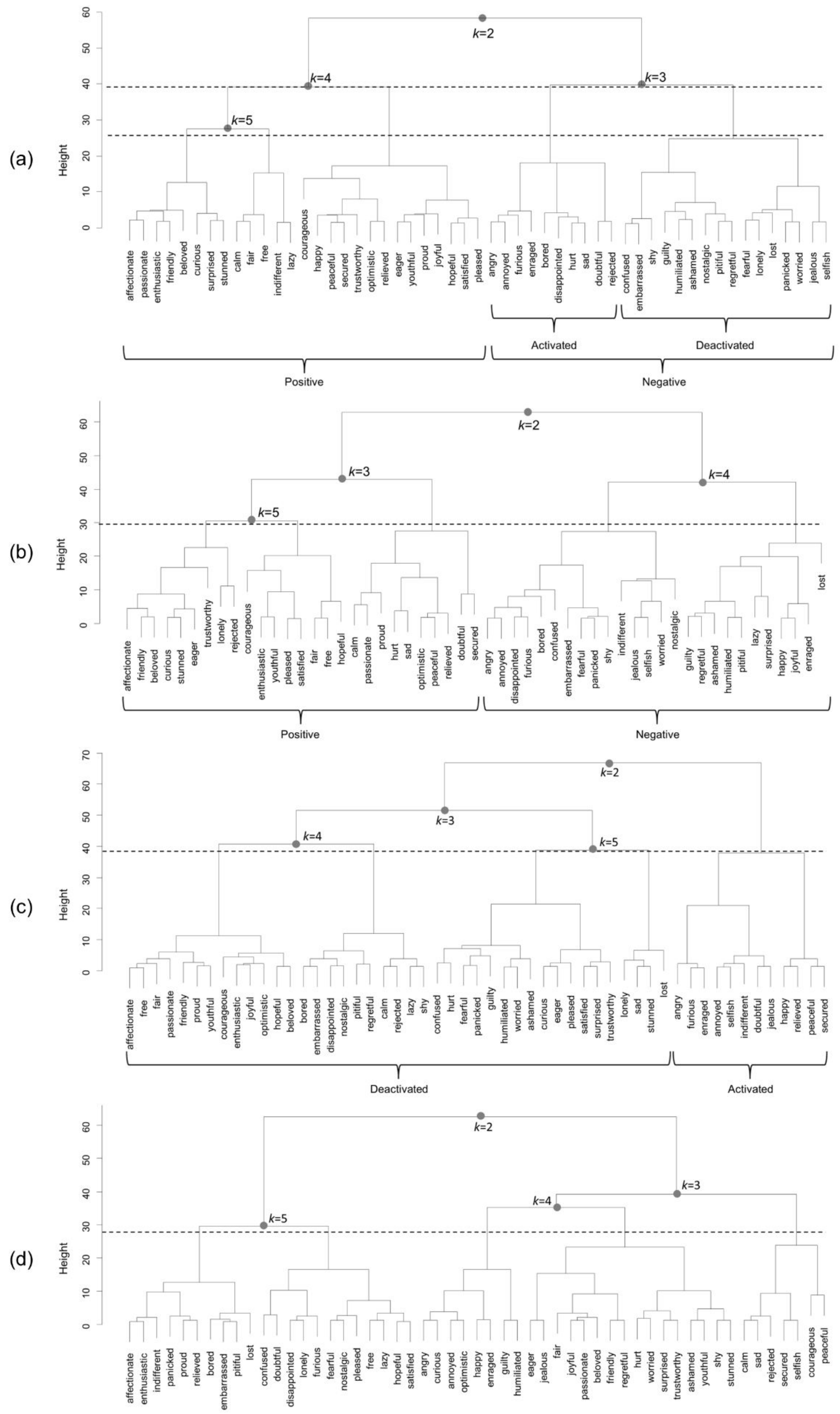


Figure 2.7 Illustration of hierarchies at subject level obtained by divisive algorithm. (a) hierarchy of subject 1, (b) hierarchy of subject 11, (c) hierarchy of subject 73, and (d) hierarchy of subject 59.

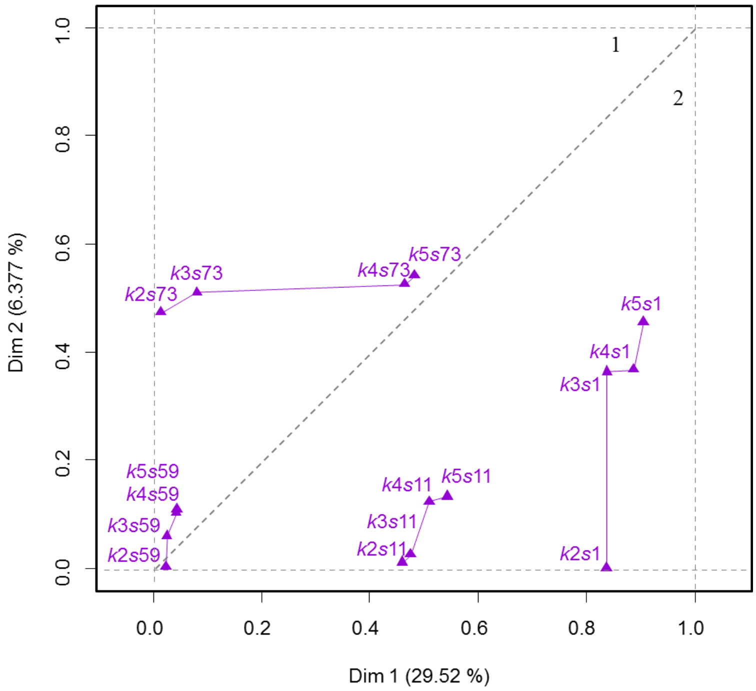


Figure 2.8 Trajectory representations of the subjects 1, 11, 59, and 73.

Remark 7: Four different cognitive process representations have been presented in this section with the idea that we can classify all subjects based on their cognitive process. Although the way of classifying the cognitive process of subjects is not included in this paper, in Figure 2.11 we present the 40 subjects (33%) who have the same cognitive process as that of subject 1; the 29 subjects (24%) who have the same cognitive process as that of subject 11; the 21 subjects (17%) who have the same cognitive process as that of subject 73; and the 31 subjects (26%) who have strategy totally different from that of the panel.

2.4 Discussion

In this section we will discuss two aspects of the Napping® based methodology. Firstly, practical issues of performing the Napping® task will be dissected and compared with other methods. Secondly, although the main objective of this paper is to present a methodology, we have also made an attempt at interpreting the typology of emotions in order to show the potential of the methodology.

2.4.1 Practical issues of performing the Napping® task

For exploring empirical models (i.e., the structure of emotions in this case) or even for validating existing models without a priori theories, the two most commonly used methods are the free sorting task and the pairwise similarity rating. Mainly because these two methods possess a common property: they avoid using references to neither emotion categories nor emotion dimensions; thus, the obtained models are not influenced by measurement methods in which scales can be theoretically or empirically oriented. The way of collecting data approached from the Napping® method that we present in this article possesses this same property.

In fact, the important information obtained by the Napping® method is the two prominent dimensions implied by each subject's configuration. From this information, we can obtain more than just the empirical model (dimensional or categorical) at a panel level, we can also comprehend subjects individually via their sequence of steps that leads to their configuration. As far as we know, the statistical analysis of this sequence of steps has never been done before. In addition, the experiments as such in this study are easy to perform since it takes on average approximately 15 minutes per subject.

Although the Napping® experiment has the advantage over the free sorting task and the pairwise similarity rating of being more informative, care must be taken that the subjects have a global view of the stimuli.

In the case of the pairwise similarity rating, the subjects take a more local view in rating similarity of each pair of stimuli. The reason being that in the case of pairwise similarity rating the subjects have to remember two stimuli only, whereas in the case of the Napping®, the subjects have to take into account the whole corpus of stimuli. It must be remembered that in spite of the fact that the pairwise similarity rating is easy to perform, there are two main limitations. Firstly, the criteria for rating similarity of one pair may change from one pair to another. Secondly, the task becomes unwieldy and inconvenient when the number of stimuli is large (e.g., 50 stimuli require 1225 comparisons).

In the case of the free sorting task, the subjects are required to have a more global view than the pairwise similarity rating as they are confronted with the whole corpus of stimuli at the same time. However, the relationship among clusters and the distances among pairs of stimulus are never asked of subjects. Hence in practice, it seems that the subjects focus on subsets of stimuli in order to ensure their homogeneity first, without thinking of an eventual relationship between subsets.

Returning to the perspective of the data collection, it seems that the Napping® method requires the subjects to remain focused throughout the task more so than the pairwise similarity and the free sorting task. Actually, the subjects have to take into account both the distances amongst these emotions as that of the pairwise similarity and amongst subsets of emotions as that of the free sorting task. As a consequence, it seems to be even more complicated when the subjects deal with a large numbers of stimuli.

The observed data shows that 121 of a total of 140 subjects (86.4%) completed their configurations. Of the 19 subjects who did not succeed in completing their configurations, we could assume that either they could not understand or did not have enough patience for the Napping® task. Still, despite this, the majority of subjects had managed to perform the task, which shows that a subject can deal with a large range of emotions (in our case more than 50 emotions). However, based on our observations, we recommend that sixty should be the maximum number of emotions in relation to the time taken and the complexity of the Napping® task. Still, it is possible to exceed this limit by using balanced incomplete block designs.

2.4.2 Interpreting emotion typology of Vietnamese culture

A number of emotion researchers has postulated that the dimensional and categorical approaches are complementary rather than contradictory. We totally agree with this concept, and we feel we can go further. In this section, we will uncover the coexistence of the dimensional and categorical approaches obtained by the Napping® methodology using a dynamic structure as presented throughout this article. In addition, we formulate an interpretation of emotion typology which is a balance between the universal human values invariant across cultures and the social construction specific to Vietnamese culture. Before starting the interpretation, we would like to stress that we consider each cluster as an emotion family in the following sense that each family consists of at least a basic emotion, some related emotions, and perhaps even characteristic emotions of the culture.

Figure 2.9 illustrates 5 clusters (cf. Table 2.2) obtained by the divisive algorithm positioned in the factorial planes constituted by the first three dimensions obtained by MFA. As can be observed, cluster 1 and cluster 2 are opposite to cluster 4 and cluster 5. This opposition constitutes the first dimension which is a universal dimension: the valence (i.e., the positive-negative or pleasant-unpleasant dimension).

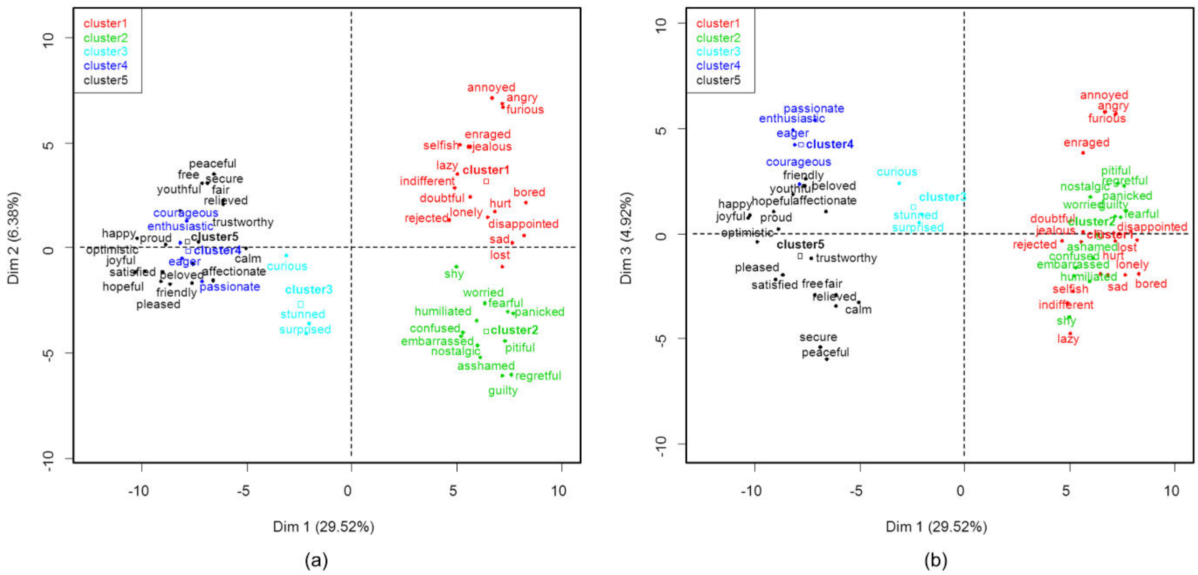


Figure 2.9 The typology of 5 clusters was illustrated on the factorial planes obtained by MFA: (a) the first-two dimensional plane, and (b) the first-three dimensional plane.

Remarkably, cluster 3 is positioned between clusters 1 & 2 on the one side and clusters 4 & 5 on the other side. As mentioned in Table 2.2, cluster 3 consists of the three emotions "curious", "surprised", and "stunned". These emotions, from our point of view, can be attributed in any valence depending on the case or event. For instance, "surprised" and "stunned" could be seen to be experienced from an unexpected event, whereas "curious" could be seen to denote an aspect of behaviour, yet they could all be caused by a positive or negative event. Thus, it can be considered as neutral.

In the next step, the heterogeneity of cluster 1 and cluster 2 can be observed in Figure 2.9.a. They are separated in the second dimension of the first factorial plane in order to form the arousal (i.e., the activate-deactivated) dimension. Like the valence dimension, this dimension is also one of the universal dimensions.

Cluster 1 (cf. Table 2.2) refers to the two primary emotions “angry” and “sad” as seen in the early studies of Ekman & Friesen (1971) and Plutchik (1980). Beyond their presence, it consists of emotions which are associated with “angry” such as “annoyed”, “furious”, and “enraged”; and “sad” such as “disappointed”, “hurt”, “lonely”, and “rejected”. The relationships among such emotions are supported in the study by Storm & Storm (1987, Table 2 & Table 4, p. 811). They all imply a violation or an angry reaction, the so-called negative-activated cluster which may be common to the universal human values.

In accordance with Vietnamese culture is the presence of emotions such as “lazy” and “bored”. They belong to this cluster since they are characteristics that violate the social expectations or norms due to the fact that “working hard” and “joining together” are characteristics that traditional culture provides to the self in an agricultural country such as Vietnam.

Cluster 2 (cf. Table 2.2) contains emotions related to “fearful” and “ashamed”. In this, “fearful” is a primary emotion, whereas “ashamed” is considered as a primary emotion only by few authors, such as Izard (1971) and Tomkins (1984). “Fearful” is associated with “worried”, “anxiety”, and “shy” (Johnson-Laird & Oatley, 1989); whereas “ashamed” is associated with “shame”, “embarrassed”, and “humiliated” (Scherer, 2005, Table 4, p. 715), or “embarrassed”, “shy”, and “humiliated” (Storm & Storm, 1987, Table 3, p. 811). Comparing this with the first negative cluster, the negative emotions of this cluster are of lesser intensity and imply a minor provocation, the so-called negative-deactivated cluster. Lack of specificity with respect to Vietnamese culture is observed in this cluster.

The two first steps are consistent with the two dimensions of the first factorial plane which form the two universal dimensions. It is difficult to observe the homogeneity of cluster 4 and cluster 5 on this plane. Therefore, the second factorial plane is needed to exploit the two positive clusters. From our point of view, this third dimension together with these two clusters is particular to Vietnamese culture.

Cluster 4 (cf. Table 2.2) consists of emotions “courageous”, “eager”, “passionate”, and “enthusiastic” which represents the “interpersonal reference” and integrates social standards. Throughout the 4000 year plus history of Vietnam, the people have been “eager”, “enthusiastic”, and “courageous” as they have dedicated themselves to the country, for example in confronting floods, drought, and invaders. These emotions honestly express the characteristics of the nationalism of Vietnamese people (Phan, 1998, p. 34-45).

Cluster 5 (cf. Table 2.2) consists of positive emotions within a narrower sphere. For instance, the emotions such as “affectionate”, “trustworthy”, and “peaceful” related to the self and the related-self which always follow the national interests. In accordance with Vietnamese culture, these emotions reflect the relationship between the self, family and the local community on a village scale in daily life, the so-called Vietnam village culture (Phan, 1998, p. 61-62).

Remark 8: Although we focus on the typology of five clusters based on its interpretation, it does not mean that this typology is a unique solution for classifying emotions. From our point of view, the number and underlying meaning of dimensions and clusters is a researcher–defined work which depends on the working fields and experience of researchers.

A Napping® based methodology to quickly obtain a model of emotions - some complementary thoughts

Chapter 2 is an occasion to introduce the heart of our research, where the concept of cognitive process associated to the Napping® task is introduced, defined, and formalised. In the framework of this chapter, the formalisation of the cognitive process is based on the subject behaviour model, which is assumed when subjects deal with a Napping® task. For that purpose, two assumptions are made: first, subjects perceive all the stimuli using a top-down strategy; and second, the top-down strategy is limited to 5 groups of stimuli. These assumptions are based on behavioural observations when subjects deal with a Napping® task.

The top-down strategy seems quite natural in a holistic context where stimuli are perceived at a more global level. Using the same analogy as the one presented in chapter 1, a photographer in this context would use a zoom lens with small focal distance to take a series of pictures, and regarding that lens would examine the stimuli from quite far away, hence the number of groups that would range from 2 to 5. The number of groups formed by a given subject can be explained mechanically by the physical constraints of the Napping® task, notably the rectangle shape on which the stimuli are positioned. Apparently, two dimensions constitute 4 poles, and consequently 4 groups on the rectangle, the 5th group corresponding to a division of one of the 4 groups that would allow the subject not to provide a too “trivial” configuration.

Once the model admitted, Figure 2.7 and Figure 2.8 suggest to automatically classify the trajectories. Let’s remind that the trajectories are the representations of the cognitive processes on the dimensions stemming from the MFA. The issue of classifying the trajectories itself is an open question relating to the determination of the distance between trajectories in a multidimensional space. By construction, a trajectory (of a given subject) can be assimilated to a series of quadruplets, each quadruplet corresponding to the coordinates of each level (of the hierarchy) on the dimensions provided by the MFA (cf. Figure 2.10).

Naïve but practice, the Euclidean distance applied to the data set as represented in Figure 2.10 can be interpreted the following way: two trajectories are close if the two subjects have a similar modelled behaviour concerning the Napping® task. Thus, by using the Euclidean distance to the data set of Figure 2.10, we can perform a classification of trajectories, i.e., the cognitive process of subjects. These results are presented in Figure 2.11.

Remark: The Euclidean distance used for the classification analysis does not take into account the nature of the data. In detail, it ignores the relative importance of one dimension with respect to other dimensions; and within one dimension, it ignores the natural orders of one column with respect to other columns (i.e., L1, L2, L3, and L4; cf. Figure 2.10).

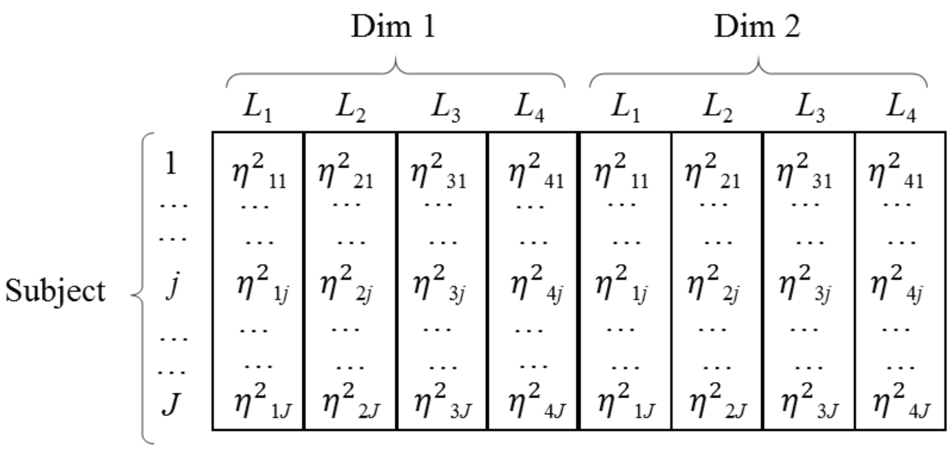


Figure 2.10 Illustration of the structure of data table for the analysis of classifying the cognitive processes of subjects using agglomerative clustering

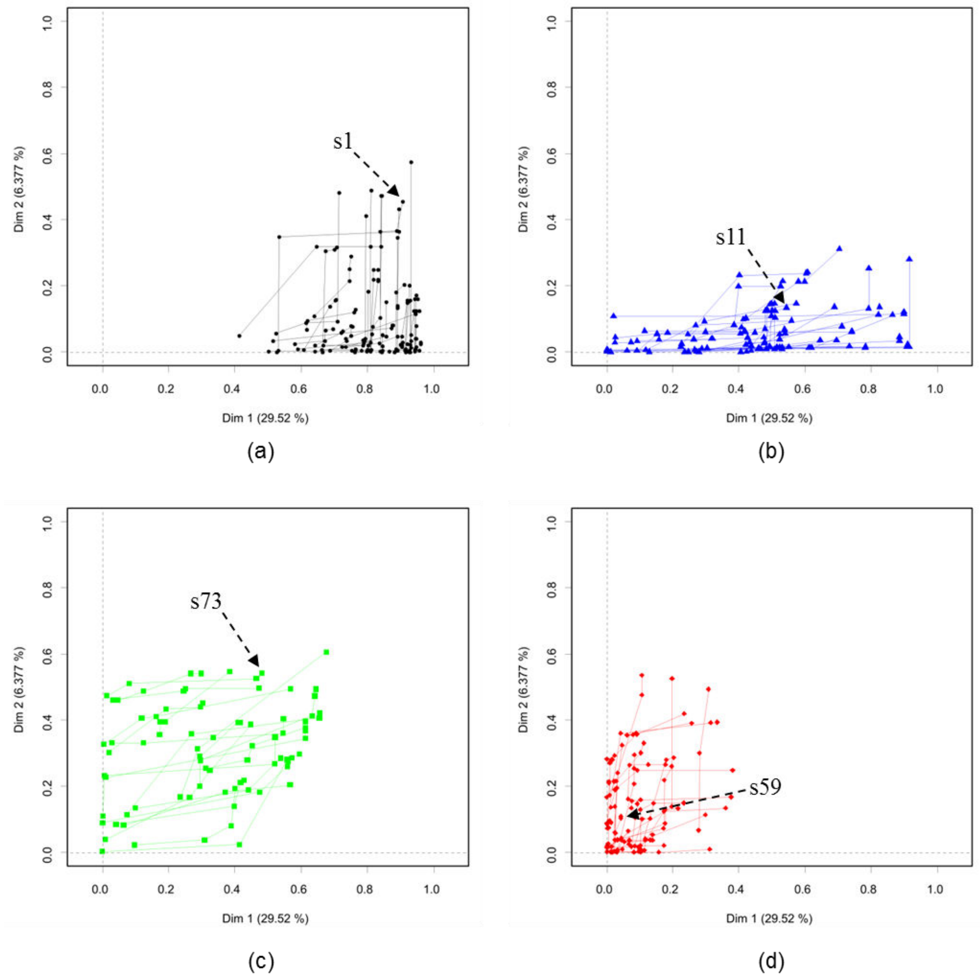


Figure 2.11 The four different cognitive processes of the subjects, who have the same strategies for clustering emotions: as subject 1 (a), as subject 11 (b), as subject 73 (c), and as subject 59 (d)

In our example, the 121 Vietnamese subjects who conducted the Napping® task on the set of emotions can be separated into 4 groups:

- Of 33% (40 subjects) who have the same strategies for clustering emotions as subject 1 (cf. Figure 2.7.a, Figure 2.8, and Figure 2.11.a)
- Of 24% (19 subjects) who have the same strategies for clustering emotions as subject 11 (cf. Figure 2.7.b, Figure 2.8, and Figure 2.11.b)
- Of 17% (21 subjects) who have the same strategies for clustering emotions as subject 73 (cf. Figure 2.7.c, Figure 2.8, and Figure 2.11.c)
- Of 26% (31 subjects) who have the same strategies for clustering emotions as subject 59 (cf. Figure 2.7.d, Figure 2.8, and Figure 2.11.d)

The representation of the cognitive process of a subject depends on the model of subject behaviour following the top down strategy. For instance, in the case of subject 1, the model does not allow to obtain a trajectory such as the one presented in Figure 2.12.a: this trajectory could be associated with subjects who perceived first the dimension activated-deactivated and then the dimension positive-negative. Likewise, for the subject 11, the divisive model does not allow to obtain a trajectory such as the one presented in Figure 2.12.b: this trajectory could be associated with subjects who first perceived the dimension activated-deactivated, then who changed their mind to position the stimuli according to the sole positive-negative dimension.

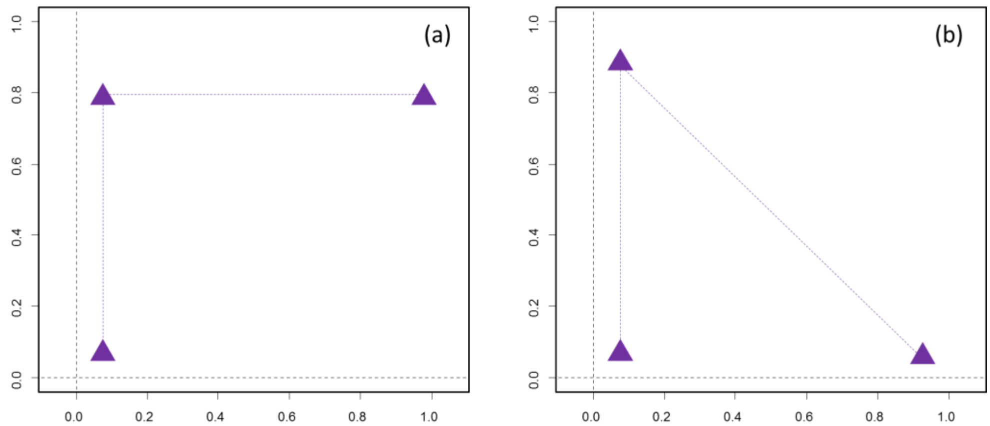


Figure 2.12 Two examples of cognitive process that cannot be modelised if the top-down approach is assumed

Figure 2.12 shows the limits of the proposed model through the two examples of cognitive processes that can not be obtained by the model. These limitations are inherent from the way the data are collected where only the final configuration provided by subjects is focused on. To overcome this shortcoming, perception of subjects during the task should be measured over time. This is precisely the objective of chapter 3, which introduces the notion of temporal Napping® data.

(This page is left blank)

3

Digit-tracking: interpreting the evolution over time of sensory dimensions of an individual product space issued from Napping® and sorted Napping

A final attempt to integrate subject's cognitive process evolving over time in analysing Napping® data

Les deux premiers chapitres de ce manuscrit nous ont permis de prendre conscience des limites de l'utilisation des données de Napping®, telles que recueillies originellement, pour comprendre le processus cognitif du sujet face à cette tâche.

Utilisée telle quelle, la configuration des stimuli fournie par les sujets, à travers la représentation des groupes en AFM (cf. Figure 1.5, chapitre 1), ne permet pas de comprendre le processus cognitif dans son ensemble, mais seulement dans sa phase finale.

Cette même configuration, adossée à une hypothèse forte sur le comportement des sujets, à savoir l'utilisation d'une stratégie descendante lors du Napping®, permet d'obtenir une représentation sous forme d'une ligne brisée des processus cognitifs (cf. Figure 2.8, chapitre 2). Le modèle proposé, en apparence séduisant théoriquement, ne semble pas opérationnel. D'une part, car il assimile un comportement "continu", celui du Napping®, à un comportement "discret" celui du tri hiérarchique; d'autre part, car le tri hiérarchique ne permet pas au sujet de modifier sa perception des stimuli au cours de la tâche.

Cette dernière limitation est extrêmement contraignante lorsque le sujet est confronté à un ensemble de stimuli complexes, pour lesquels la perception peut évoluer soit d'une évaluation à l'autre, soit au cours du temps.

La Figure 3.1 illustre le changement de perception d'une évaluation à l'autre. Suivant l'angle de vue utilisé, un même sujet peut percevoir le cube de Necker soit avec une face avant de couleur jaune, soit avec une face arrière de couleur jaune.

Au-delà de cet exemple, qui peut paraître anecdotique, des méthodes de recueil de données sensorielles ont été développées afin de comprendre les évolutions de perception au cours du temps. On peut citer par exemple la méthode TDS, pour “Temporal Dominance of Sensations”, qui permet d’identifier les descripteurs sensoriels perçus comme dominants au cours du temps (Pineau & Schlich, 2015).

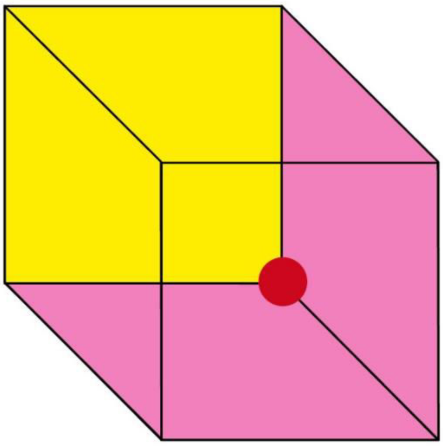


Figure 3.1 A Necker cube – An example of perceptual shifting phenomenon on human vision (Available at https://www.at-bristol.org.uk/neckercube.html)

La Figure 3.2 extraite de Pineau & Schlich (2015) illustre le fonctionnement de la méthode TDS. Le sujet sélectionne un descripteur sensoriel parmi ceux proposés, au cours du temps, chaque fois qu’il le juge dominant.

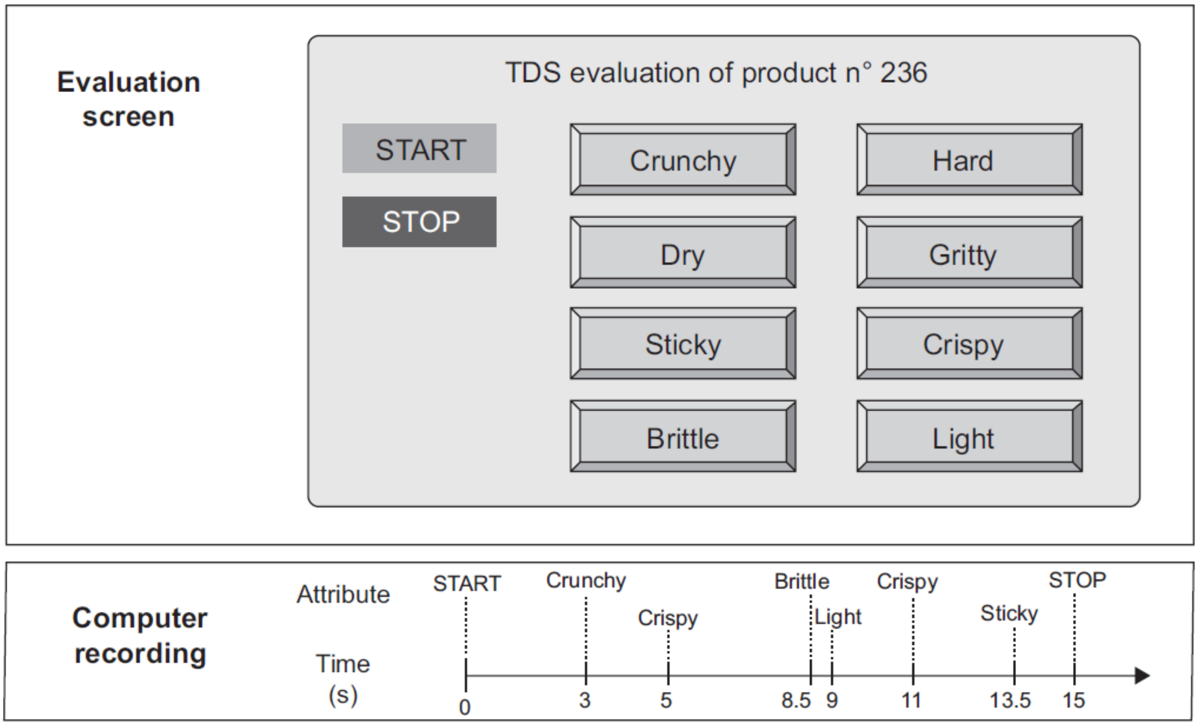


Figure 3.2 An example of TDS method on cereal product, which records subject’s perception evolving over time during a tasking period (Pineau & Schlich, 2015, p. 270)

L'objet de ce chapitre 3 est de présenter une nouvelle méthode de recueil de données qui permet de mesurer l'évolution des distances entre les stimuli au cours du temps, lors de la tâche de Napping®. Cette méthode permet non seulement de mieux comprendre le processus cognitif des sujets, mais également de mieux comprendre les dimensions sensorielles de l'espace produit issu de l'ensemble des sujets.

Ce chapitre correspond à l'article publié dans le cadre du numéro spécial *Food Quality and Preference*, qui fait suite au congrès Sensometrics 2014 (Chicago, IL, USA).

Les données utilisées sont extraites d'une expérimentation dont l'objectif principal était de comprendre l'adéquation entre la perception d'un ensemble de parfums et leur campagne publicitaire respective. Les sujets ont donc été amenés à évaluer à la fois des parfums et des vidéos, à l'aide de l'environnement Holos également présenté dans ce chapitre, mais sur lequel nous reviendrons plus en détail dans le chapitre suivant.

Chapter outline

This chapter includes the publication:

Lê, M.T., Husson, F., & Lê, S. (2015, accepted). Digit-tracking: interpreting the evolution over time of sensory dimensions of an individual product space issued from Napping® and sorted Napping, Food Quality and Preference. Doi: 10.1016/j.foodqual.2015.07.002

Abstract

Up to now, the sole information used when analysing (sorted) Napping® data has been the coordinates of the stimuli on a plane (and the way they have been grouped). This information corresponds to the final configuration provided by one subject. Beyond this final configuration, it is interesting to observe what the subject actually did during the whole experiment.

Rather than focusing on the final configuration only, our starting point is to collect and analyse data from the moment the stimuli are presented, to their final configuration. The rationale behind is to explore the evolution over time of the sensory dimensions that structure a set of stimuli.

In this paper, we propose a new methodology for studying the evolution over time of sensory dimensions of an individual product space issued from Napping® or sorted Napping. This methodology is based on a technique for collecting data, named digit-tracking: data are collected on a tactile device, and stored on a server that depends on the collaborative environment Holos, specifically developed for that purpose. The proposed methodology is presented through a case study on advertising videos of ten luxury men perfumes.

From a product perspective, our methodology helps in assessing the relative importance of one dimension on another dimension. From a subject perspective, it helps in understanding the different individual cognitive processes (i.e., the different behaviours) and in differentiating them.

Keywords: digit-tracking, evolution over time of sensory dimensions, individual product space, (sorted) Napping®, Holos environment.

3.1 Introduction

In sensory analysis, one of the main reasons to ask subjects to assess stimuli is to comprehend the sensory dimensions that discriminate among these stimuli. Reductively, the sensory comprehension of a set of stimuli can be regarded as the identification of the sensory dimensions that best separate the stimuli; in other words, the dimensions that structure the set of stimuli. The way these dimensions can be revealed has evolved over time, along with the techniques used to assess the stimuli.

In a very simplified form, Quantitative Descriptive Analysis (QDA®; Stone et al., 1974) starts from defined sensory attributes common to all subjects, to end with sensory dimensions. These dimensions are based on the best combinations of the original attributes. The practical difficulty when interpreting the sensory dimensions lies in the interpretation of the coefficients of the combinations of the original attributes: what are the latent variables behind these combinations of attributes?

Free Choice Profiling (FCP; Williams & Langron, 1984) starts from individually elicited attributes to also end with sensory dimensions based on linear combinations of attributes. The difficulty when interpreting the results obtained from FCP is compounded compared to QDA® as the attributes that are analysed are individually elicited: homology among attributes that have the same name may not be that obvious when these attributes are used by different subjects.

Finally, we could say that holistic methods such as Napping® or sorting start directly from individual sensory dimensions to end with sensory dimensions based on combinations of individual sensory dimensions. Expressed this way, the difficulty in comprehending the sensory dimensions is to be multiplied manifold compared to QDA® or FCP: as subjects are given more and more freedom in the way they assess stimuli, the resulting sensory dimensions are getting more and more difficult to interpret. To overcome this difficulty, subjects are often asked to describe the stimuli individually or by group. Typically, this information is invaluable as it allows to complete the analyst's own expertise on the stimuli.

Beyond their meaning, another way of comprehending the dimensions is to assess their stability. In the case of Napping® for instance (Pagès, 2005), the sensory dimensions are first understood by the analysis of the representation of the stimuli provided by MFA. Then, the dimensions are understood by the analysis of the representation of the subjects through the so-called representation of the groups of variables in MFA: in this representation, two subjects are all the more close as their perceptions of the set of stimuli are close. Hence, the more consensual the subjects, the more stable the dimension.

The aim of the paper is to present a new point of view and a new methodology that complete the analysis of the stability of the sensory dimensions, in the framework of Napping® and sorted Napping. This new point of view goes beyond the study of the individual sensory dimensions provided by the subjects as it integrates their evolution throughout time. In certain contexts, the process of doing a certain task can be assimilated to the cognitive process of the subject performing the task. The point of view introduced in the paper is based on the study of the individual cognitive processes. We expect that studying the successive steps that lead a subject to their final configuration might aid in understanding the notion of consensus amongst subjects: a sensory dimension has to be considered differently whether subjects have revealed it in the same way throughout time or in many different ways. Though all roads may lead to Rome, did the subjects use the same cognitive process or did they use different ones?

Preliminary work on the cognitive process of subjects in a holistic method context was done by Cadoret et al. (2011). In their paper, the authors proposed a methodology to represent the steps of the cognitive process of subjects when performing a hierarchical sorting task. This task can be seen as a succession of sorting tasks. As such, the subject provides not only a partition on the stimuli, but a sequence of nested partitions. Cadoret et al. (2009) proposed to plot this sequence of nested partitions by first plotting each partition in the so-called representation of the groups of variables provided by MFA, then by connecting these partitions, through a line, chronologically. This representation revealed that sensory dimensions that were common to most subjects could be classified into two groups: the ones that are attained by using one sensory way, by opposition to the ones attained using at least two ways. Inspired by Cadoret et al. (2011), we propose a new methodology for studying the cognitive process of subjects when performing Napping® or sorted Napping. In other words, we investigate the evolution of the sensory dimensions that structure each individual perceptual space over time. This methodology consists of a new technique for collecting data, and a new way of visualizing these data in order to analyse them and interpret them.

The following section consists of three main parts. The first part presents a new technique for collecting temporal data issued from Napping®. This technique is called digit-tracking, and can be used thanks to the Holos environment, developed for that particular purpose. The second part presents a new way of visualizing digit-tracking data in order to analyse them and interpret them, through graphical outputs provided by the new functions integrated in the R package SensoMineR, and developed for that particular purpose. Finally, the third part presents a case study on advertising videos of ten luxury men perfumes in order to demonstrate the proposed methodology.

3.2 Methodology

3.2.1 Digit-tracking: a new technique for collecting temporal data for Napping® and sorted Napping

To approach the cognitive process of a subject throughout the duration of the experiment rather than its end, the idea is to record chronologically all the steps that lead to a subject's final configuration: the final configuration, in the case of Napping®, is defined by the relative position of the stimuli on a plane of dimensions 60 by 40 centimetres.

To do so, the experiment can be conducted by using the Holos environment (Lê & Lê, 2014). Holos is a new collaborative environment in which researchers can carry out experiments based on holistic methods, such as sorting, Napping®, and sorted Napping, on various types of stimuli, such as texts, images, sounds, and videos. With Holos, researchers can also share their study resources (stimuli and data collected), either partially or totally, within the scientific community.

Technically, Holos can be defined as a combination of an Android application (version 4.4; Google Inc., 2014) for collecting data and a database for storing data. The application works based on a tactile human machine interface, which is a graphical interface that allows a subject to touch interaction with a tablet device. The database is a web server for storing all the data collected. The Holos application can be downloaded at the following address: http://napping.agrocampus-ouest.fr/.

In Holos, all the stimuli are presented randomly on the very left side of the tablet. They are displayed as icons that can be moved at will. The experimenter chooses the icons *via* the Holos interface. When performing a Napping® or a Sorting, double tapping an icon zooms the subject in when the icon represents an image or a text, or plays the sound or the video when the icon represents a sound or a video. As the subject is moving the stimuli on the tablet, their coordinates are recorded. More precisely, for a given subject, the application is generating a matrix of dimensions $I \times 2$ each time a stimulus has been moved beyond a radius of 3 pixels, where I is the number of stimuli and 2 the number of coordinates recorded (i.e., the X-axis and the Y-axis). Afterwards, we will use the term *step* each time a stimulus has been moved (i.e., each time we generate a new matrix). By recording all these coordinates, it is possible to reconstitute the trajectories of all the stimuli - and hence, to analyse the evolution of the sensory dimensions of a subject, by looking at the evolution of the relative positions of the stimuli throughout time.

By definition, a digit is any number from 0 to 9, or a finger – tracking is the process of following something or someone – the technique was named digit-tracking. By analogy with eye-tracking, digit-tracking is the process of tracking the motion of a digit (i.e., the coordinates of the stimuli) when a subject performs Napping® or sorted Napping.

Figure 3.3 illustrates some of the 40 successive steps used by subject 42 when performing a Napping® task on a tablet. In this experiment, stimuli are displayed on the screen and are dragged using a finger. The subject is asked to position the stimuli on the screen in such a way that two stimuli are close if they were perceived as similar and distant from one another if they are perceived as different.

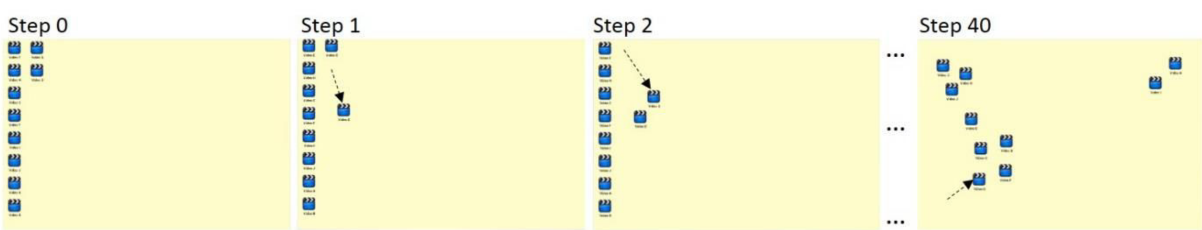


Figure 3.3 Representation of the different steps taken by a subject when performing a Napping® task on a tablet device. In this example, it took 40 steps for subject 42 to complete the Napping® task.

To be consistent with our main objective, which is the study of the individual sensory dimensions, data can be seen slightly differently. For a given step *t* and a given subject *j*, the matrix of the coordinates of the stimuli $X_j(t)$ can be transformed into a scalar product matrix among the stimuli, which contains the relative positions among the stimuli. This scalar product matrix among the stimuli, denoted $W_j(t)$, is of utmost importance as it is the object represented in the so-called representation of the groups of variables in MFA (Pagès, 2015). The information contained in this matrix can be interpreted as the shape of the set of stimuli as perceived by subject *j* at time *t*. In other words, it represents the main opposition between the products, i.e. the sensory dimensions of subject *j* at time *t*.

3.2.2 *Analysing digit-tracking data*

As in Cadoret et al. (2011), the analysis of the individual sensory dimensions along with their evolution is done with respect to the consensual final configurations of the stimuli. In other words, the representation of an individual cognitive process is obtained by projecting their $W_j(t)$ as supplementary elements on the axes induced by the MFA performed on the individual final configuration. This way of representing and analysing the individual cognitive processes answers the following question: how have the sensory dimensions of a given subject evolved throughout time with respect to the final configuration based on all subjects?

As a consequence, the analysis of digit-tracking data is performed in two important steps. The first one consists in performing an MFA on the individual final configurations in order to obtain a reference framework: this reference framework is the consensual representation of the set of stimuli, in which the part of each subject has been balanced. The second one consists in projecting the $W_j(t)$, as supplementary elements, on the axes of the representation of the groups of variables provided by MFA.

Let's recall the following properties on the projection of the $W_j(t)$ as supplementary elements. The coordinates of the $W_j(t)$ (or equivalently of the $X_j(t)$ matrix) on the components $(F_s, s \geq 1)$ induced by the MFA on the final individual configurations is defined as the *Lg* measure between the matrix $X_j(t)$ and the components $(F_s, s \geq 1)$. For instance, in the case of Napping® data, as $X_j(t)$ is constituted of continuous variables, the coordinate of $W_j(t)$ on the axis of rank *s* is equal to $Lg(F_s, X_j(t)) = \frac{1}{\lambda_1^j} \sum_{k \in K_j} r^2(F_s, X_j(t))$, where λ_1^j is the first eigenvalue of the unstandardized PCA on $X_j(t)$. The *Lg* measure lies between 0 and 1: it is equal to 0 if none of the variables of $X_j(t)$ is correlated with F_s, and it is equal to 1 if F_s is the first principal component of $X_j(t)$.

3.2.3 A case study

The data used to illustrate our methodology are extracted from a study which aim was to understand, for a given set of products, the link between sensory characteristics and responses elicited to their respective advertising videos (Jourdan, Morvant, Lê, & Pagès, 2014). In this study, the products are ten luxury men perfumes: Eau de toilette Lacoste (Lacoste, London, United Kingdom), Bleu de Chanel (Chanel, Paris, France), Acqua di Gio (Giorgio Armani, Asnieres, France), One Million (Paco Rabanne, Neuilly sur Seine, France), Fleur du Male (Jean Paul Gaultier, Paris, France), Eau Sauvage (Parfums Christian Dior, Paris, France), Bottled Night (Hugo Boss, London, United Kingdom), Guerlain Homme (Guerlain, Paris, France), Ck One (Calvin Klein Cosmetics, Paris, France), and Fahrenheit (Parfums Christian Dior, Paris, France). The advertising videos used for the experiment were publically broadcasted on French television stations in 2013. The experiment was conducted at Agrocampus Ouest (Rennes, France) with 80 subjects. In this experiment, subjects had to assess in two separate sessions, the perfumes and their advertising videos. In this paper, we will limit ourselves to the Napping® data related to the advertising videos: as data were recorded using Holos, beyond the coordinates of the videos and their verbal descriptions, we also have at our disposal their trajectories.

3.3 Results

As mentioned previously, the reference space in which the cognitive processes are studied is based on the final configuration of the stimuli obtained by an MFA, applied on the final representation of each subject. Hence, at the first step, we will present this final configuration and its interpretation, at the second step, we will present the representation of the cognitive processes. In this paper, we will limit ourselves to the first two dimensions provided by MFA.

3.3.1 A preliminary analysis for understanding the representations obtained by all subjects

From a product perspective, the so-called representation of the individuals provided by MFA (cf. Figure 3.4.a) suggests a division of the 10 advertising videos into three groups. The first dimension opposes the advertising videos of Eau de toilette Lacoste and Ck One to those of Acqua di Gio, Fahrenheit, Guerlain Homme, and Fleur du Male. The second dimension opposes the videos of Bottled Night, Eau Sauvage, Bleu de Chanel, and One Million to the others.

From a subject perspective, the so-called representation of the groups of variables provided by MFA (cf. Figure 3.4.b) suggests a division of the subjects into three groups as well: (1) a first group of subjects located along the first dimension, with rather high coordinates on the first dimension and rather low coordinates on the second dimension, (2) a second group of subjects located along the second dimension, with rather high coordinates on the second dimension and rather low coordinates on the first dimension, (3) and finally a third group of subjects with low coordinates on the first and second dimensions.

The joint interpretation of these two figures expresses the fact that the first group of subjects mainly opposed the advertising videos of Eau de toilette Lacoste and Ck One to those of Acqua di Gio, Fahrenheit, Guerlain Homme, and Fleur du Male; and that the second group of subjects mainly opposed the videos of Bottled Night, Eau Sauvage, Bleu de Chanel, and One Million to the others. If we now use the verbal description of the subjects on the videos, the first group of subjects has mainly perceived the dimension driven by the concepts young - party versus nature - wild, while the second group of subjects has mainly perceived the dimension driven by the concepts seduction versus the others.

These two dimensions seem to be close in terms of variability as the first dimension expresses 25.88% of the total variability of the data, and the second dimension explains 22.39% of the total variability of the data. This result is corroborated by the distribution of our two former groups of subjects.

When interpreting the relative importance of one dimension over the other one, what happens now if we focus on the cognitive processes of each subject?

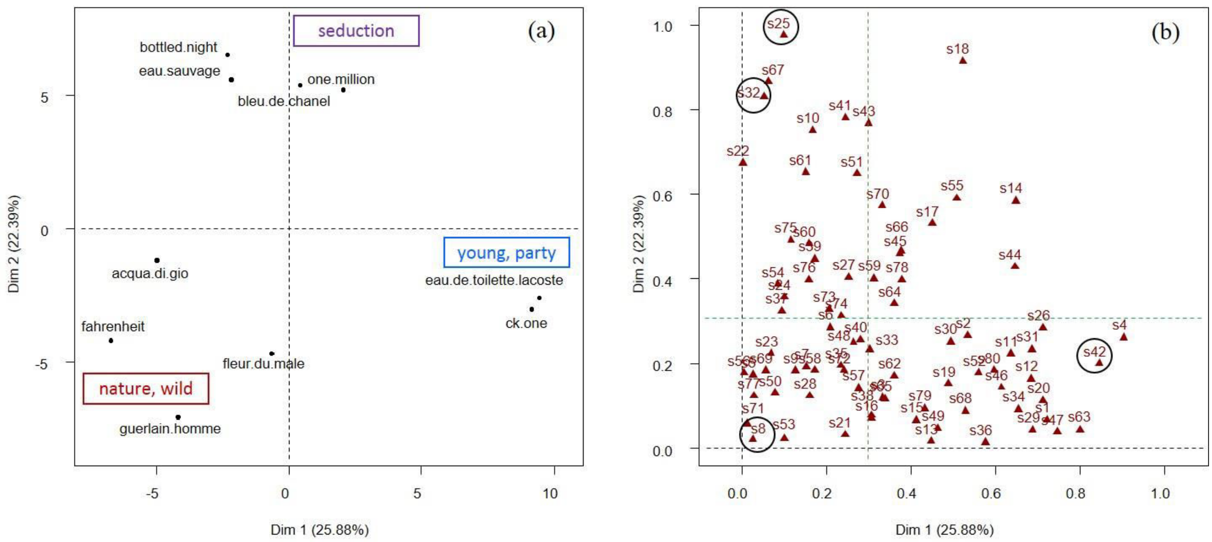


Figure 3.4 Representation of the stimuli based on Napping® data from the 80 subjects (a). Representation of the 80 subjects (b) in the so-called representation of the groups of variables provided by MFA.

3.3.2 Understanding the representations obtained by all subjects by taking into account the evolution of the sensory dimensions per subject

In this part, we propose two ways of representing the cognitive process of a given subject, in other words the way their sensory dimensions have evolved. First, each cognitive process is represented in the same space as the one used for the representation of the groups of variables provided by MFA (cf. Figure 3.5). Second, the cognitive process is represented regarding each dimension stemming from MFA; in that sense this representation tends to be interpreted as the evolution of the sensory dimensions of a subject (cf. Figure 3.6).

Figure 3 represents the cognitive processes of 4 subjects within the representation of the groups of variables provided by MFA.

Subject 8 belongs to the group of subjects with low coordinates on the first and second dimensions (cf. Figure 3.4.b). During the whole experiment, this subject has prioritized differences in dimensions other than dimensions 1 and 2 (cf. Figure 3.5.a).

Subject 42 belongs to the group of subjects with rather high coordinates on the first dimension and rather low coordinates on the second dimension (cf. Figure 3.4.b). During the whole experiment, this subject has considered the first dimension as the most important one (cf. Figure 3.5.b), i.e. the one driven by the concepts young - party versus nature - wild.

Subjects 32 and 25 belong to the group of subjects with rather low coordinates on the first dimension and rather high coordinates on the second dimension (cf. Figure 3.4.b). During the experiment, while subject 32 has considered the second dimension as the most important one (cf. Figure 3.5.b), subject 25 has first considered the first dimension as the most important one then has changed their mind and has decided to separate the videos according to the second dimension. In other words, the main sensory dimension of variability for subject 25 has evolved throughout time, which is not the case for the other subjects. The apparent proximity between subjects 32 and 25 based on their final configuration has to be tempered by their differences in terms of behaviour.

In this experiment, the representations of the cognitive processes showed that the subjects similar to subject 42 in terms of final configuration, behaved also similarly, as they did not change their mind during the experiment; the subjects similar to subject 25 represented approximately 20% of the subjects with rather low coordinates on the first dimension and rather high coordinates on the second dimension.

The fact that some subjects have behaved as subject 25 lowers the importance of dimension 2 and strengthens the importance and dominant position of dimension 1 in the interpretation. The rationale behind this re-evaluation is that a dimension is all the more strong that it has been considered as important in an unambiguous manner. Apparently, this is not the case for the second dimension for subjects who behave as subject 25.

Figure 3.6 represents the cognitive process of our 4 types of subjects, with respect to the two first dimensions stemming from MFA. The rules of interpretation of this figure are the same as for Figure 3: when the coordinate of a process on a dimension is close to one, the dimension corresponds to an important axis of variability for the process. Once the dimensions are interpreted, this representation allows the perception of the stimuli for each subject to be tracked over time, in other words the evolution of their sensory dimensions.

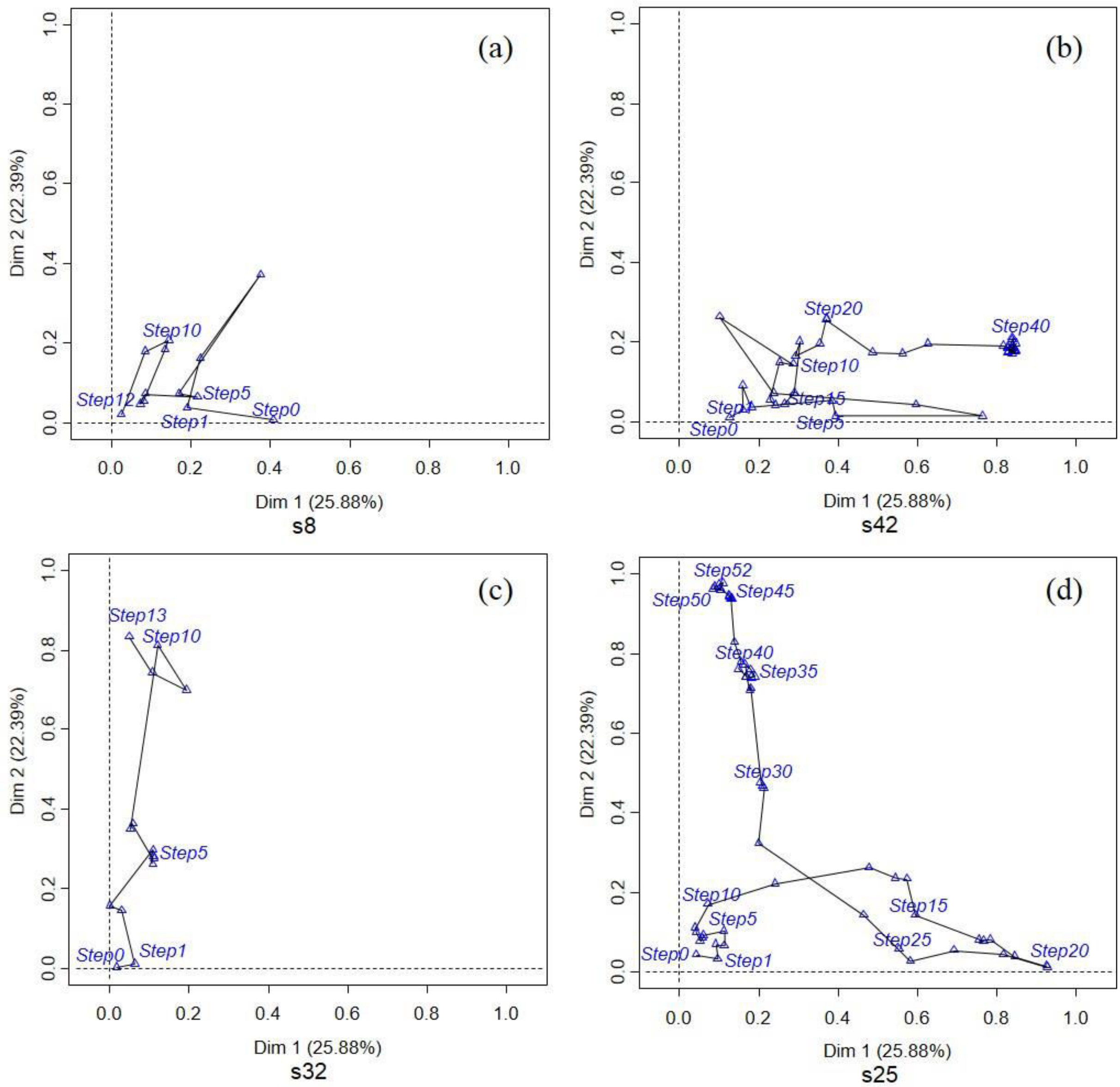
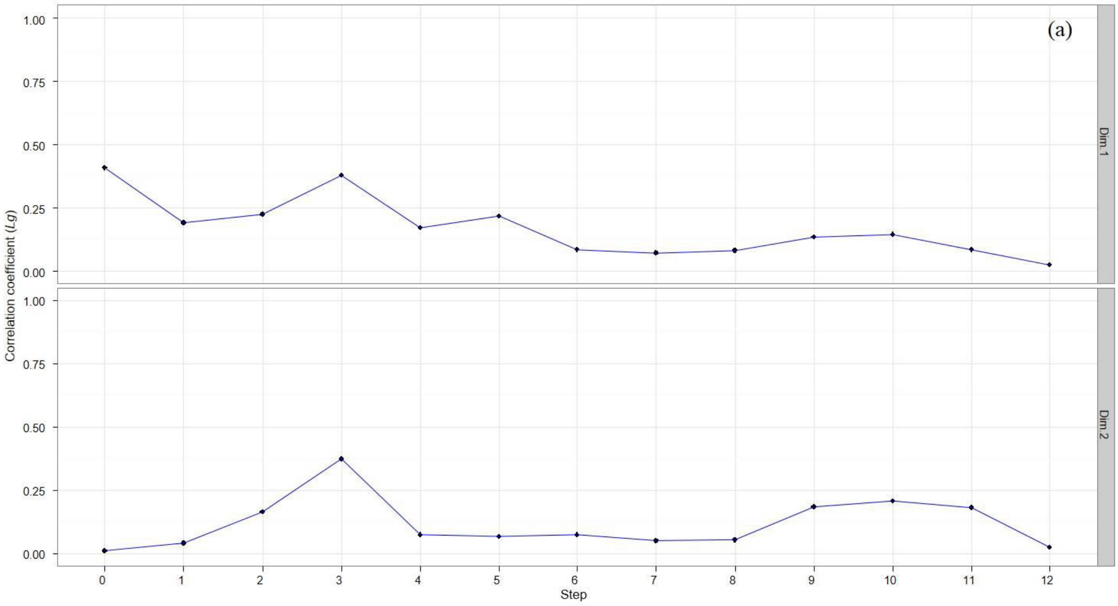
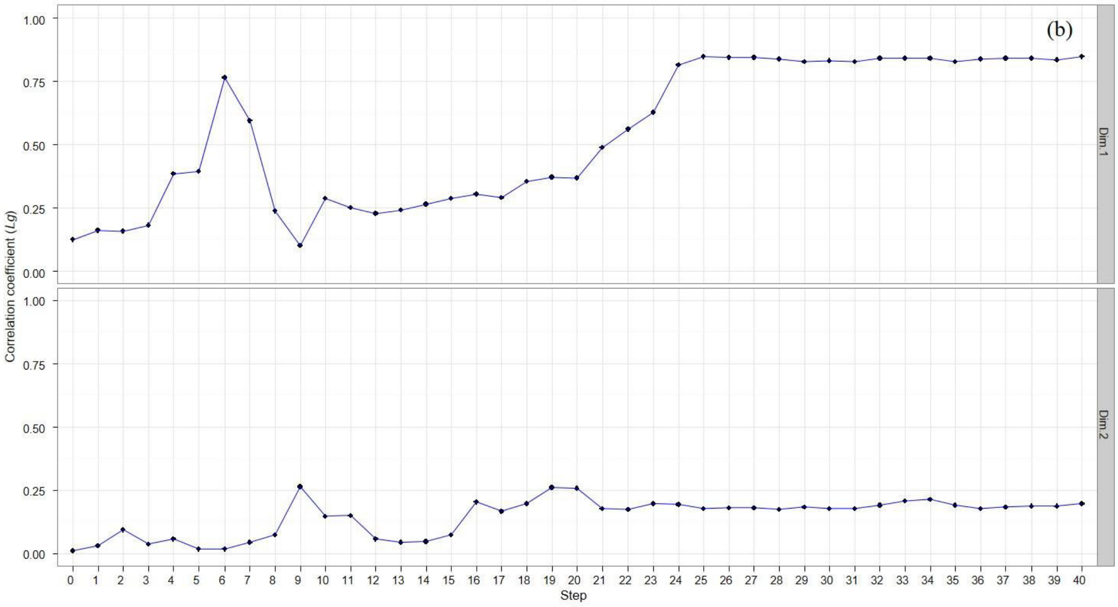


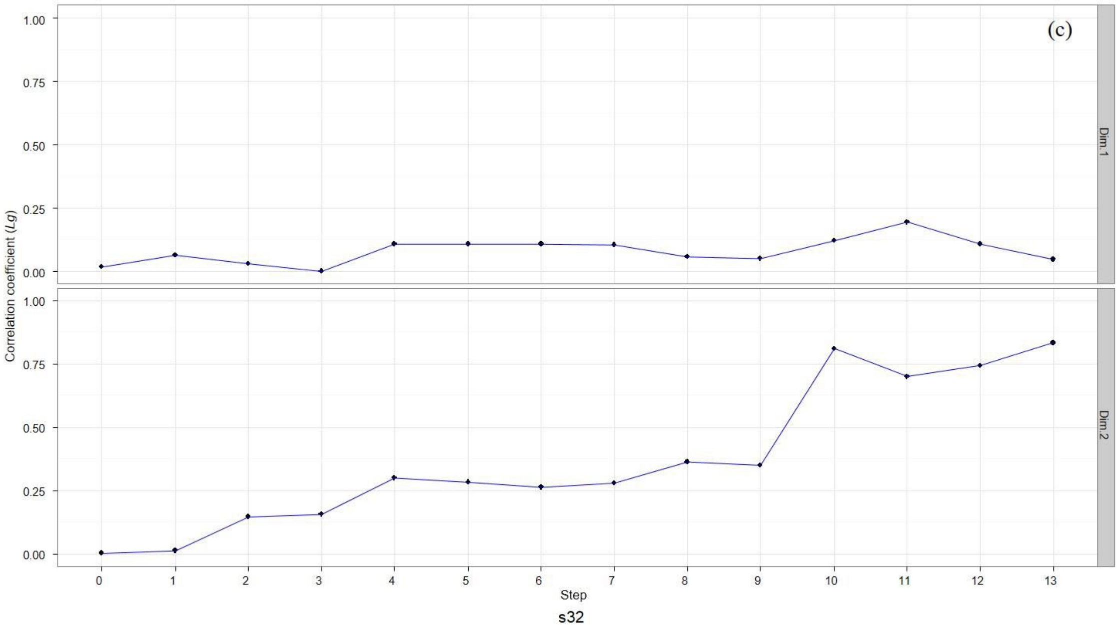
Figure 3.5 Different cognitive processes used by subjects are illustrated using results from four subjects. The steps taken by a subject a linked chronologically within the so-called representation of the groups of variables, which is provided by MFA.


(a)
Dim.1
Dim.2
Correlation coefficient (Lg)
1.00
0.75
0.50
0.25
0.00
0
1
2
3
4
5
6
7
8
9
10
11
12
Step

s8


(b)
Dim.1
Dim.2
Correlation coefficient (Lg)
1.00
0.75
0.50
0.25
0.00
Step

s42



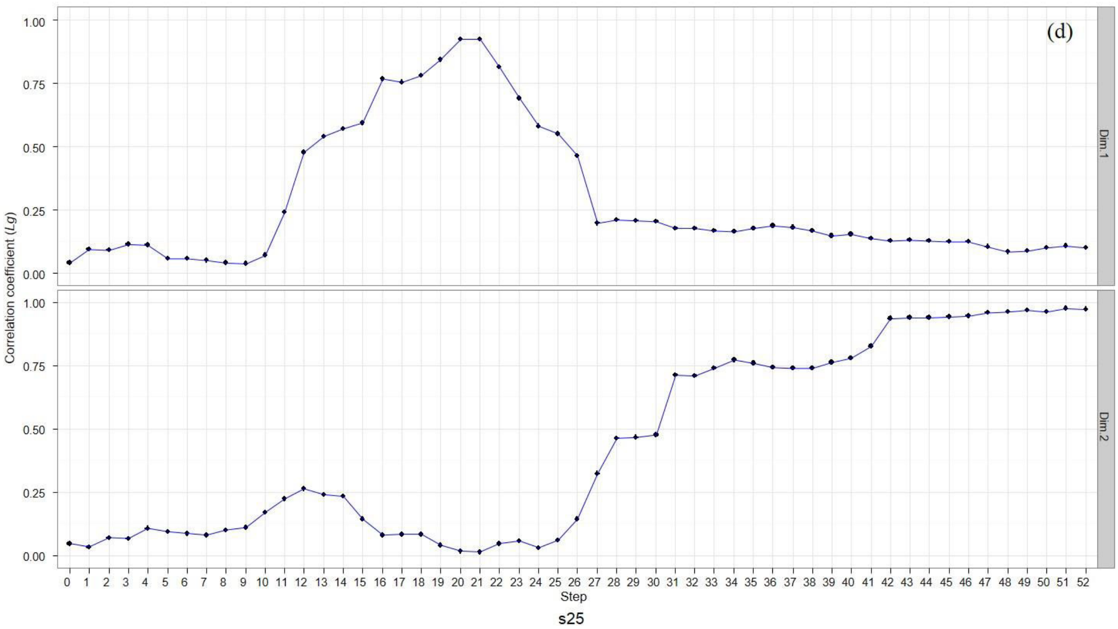


Figure 3.6 Different cognitive processes used by subjects are tracked, showing their correlation between the product configuration at each step with MFA dimensions 1 (top) and 2 (bottom), which are based only on the final Napping® data

3.4 Discussion

Beyond the results of this experiment, a questionnaire on subject satisfaction was developed to assess the use of a tablet device when performing Napping®. According to their informal feedback, subjects felt comfortable with the Holos environment, and found that performing the task on a tablet was intuitive, easy, and entertaining.

From a results perspective, subjects have clearly identified two dimensions when watching the 10 advertising videos. When analysing only final Napping® configurations, it appears that neither of the two dimensions can be considered as dominant: the variances associated with the dimensions are comparable.

This result has to be tempered with the analysis of the individual cognitive processes proposed in our paper. In this experiment, the study of the cognitive processes shows that the young - party versus nature - wild dimension has also been considered important by subjects for whom the seduction dimension was essential: this observation reinforces the dominant position of the first dimension over the second one.

From a methodological perspective, the point of view we propose in this paper to study the cognitive processes can be easily generalised to sorted Napping: in this case, the information brought by the Napping® is only reinforced by the information brought by the sorting (Lê, Lê, & Cadoret, 2015).

In terms of interpretation, work remains to be done to integrate information such as the time spent per subject, the number of steps per subject, and in general the phase during which a subject discovers the whole stimuli (learning phase).

In terms of statistical analysis, a significant work remains to be done to classify subjects according to their cognitive process: as a cognitive process can be defined as a set of curves on multiple dimensions (cf. Figure 3.6), a curves clustering algorithm that takes into account the relative importance of each dimension is under development.

Acknowledgment

We would like to thank the two anonymous reviewers for the useful comments and suggestions that help in improving the manuscript. We would like to thank the Sensometric Society for their Travel Award, which gave us the opportunity to present this work at Sensometrics 2014 in Chicago, Illinois, USA. Our special thanks goes to Dzung Hoang Nguyen, Professor at Ho Chi Minh City University of Technology (Vietnam), for supporting the Holos project.

Digit-tracking: Interpreting the evolution over time of sensory dimensions of an individual product space issued from Napping® and sorted Napping - some complementary thoughts

The previous section presented the main objective of digit-tracking, which is to study the evolution of sensory dimensions over time. But besides that, this technique can also help us to clarify two doubts (i.e., questions) relating to the way subjects conducted the Napping® task:

- Did subjects conduct the task seriously?
- At which moment subjects can discover all dimensions of a set of stimuli before making judgments on the most important dimension?

Before answering these two questions as well as sharing our complementary thoughts, it is worth noting that when subjects conduct the task, their "behaviours" are recorded and stored on the Holos server. By assessing this data, we have information relating to: (1) the time-taken, (2) the step where all the stimuli are discovered (the so-called learning step), (3) the total number of steps, and (4) the frequency each stimulus was assessed that a subject passed to complete the task.

Depending on particular purposes, researchers can be interested in one, some or all the parameters. In our context, the two mentioned questions can be answered based on the time-taken and learning step.

Did subjects conduct the task seriously?

To answer this question, we try to "detect" the subjects who might not conduct the experiment seriously based on their time-taken during the experiment.

Let's remind that in this experiment the stimuli are 10 advertising videos. Their total length are 334 seconds (around 5÷6 minutes); and each video length from 20 to 40 seconds. Therefore, in order to complete the task, it should take at least 6 minutes (here, only the main session of the Napping® task is taken into account, not including the interview, instruction, and verbalisation task).

Table 3.1 presents the time-taken for (some) subjects to complete the task. We "detect" the subjects 23 and 59, who took around 1 minutes for the task. In spite of the fact that such videos were publically broadcasted on television, they might all know, and it did not need to view the whole length of each video. But from our point of view, they might not conduct the experiment seriously as they did not assess all the stimuli (only moved and did not view all the videos).

We found that 6 subjects took less than 3 minutes to complete the task. But still, on average, the subjects took about 8-12 minutes to complete the task, and 3 subjects took much more time than the others (more than 20 minutes). It shows evidence that our subjects were serious in conducting the Napping® task on tablet device.

Table 3.1 Further information relating to subject's behaviour during Napping® task

```
> res.sum$Step.Time
 Subject Number.step Step.learning Time.taken
       8          12            12   00:02:42
      23          13            12   00:01:15
      25          52            17   00:13:25
      32          13            11   00:06:50
      42          40            13   00:08:00
      59          15            14   00:00:36
      63          39            24   00:24:11
      78          22            22   00:23:36
     ...         ...           ...        ...
```

At which moment subjects can discover all dimensions of a set of stimuli before making judgments on the most important dimension?

When conducting a qualitative observation on a Napping® experiment with three-dimensional stimuli (8 symbols), Nestrud & Lawless (2011, p.1275) observed that *"...the criteria that participants used also were not clear to them until they got two or three stimuli on the paper. Most of participant had an "aha" moment when they realised that it was impossible to easily resolve all three dimensions at once"*. This idea leads us to the question: *At which moment subjects can discover all dimensions of a set of the stimuli, and thinking about them before making judgments on the most important dimension?*

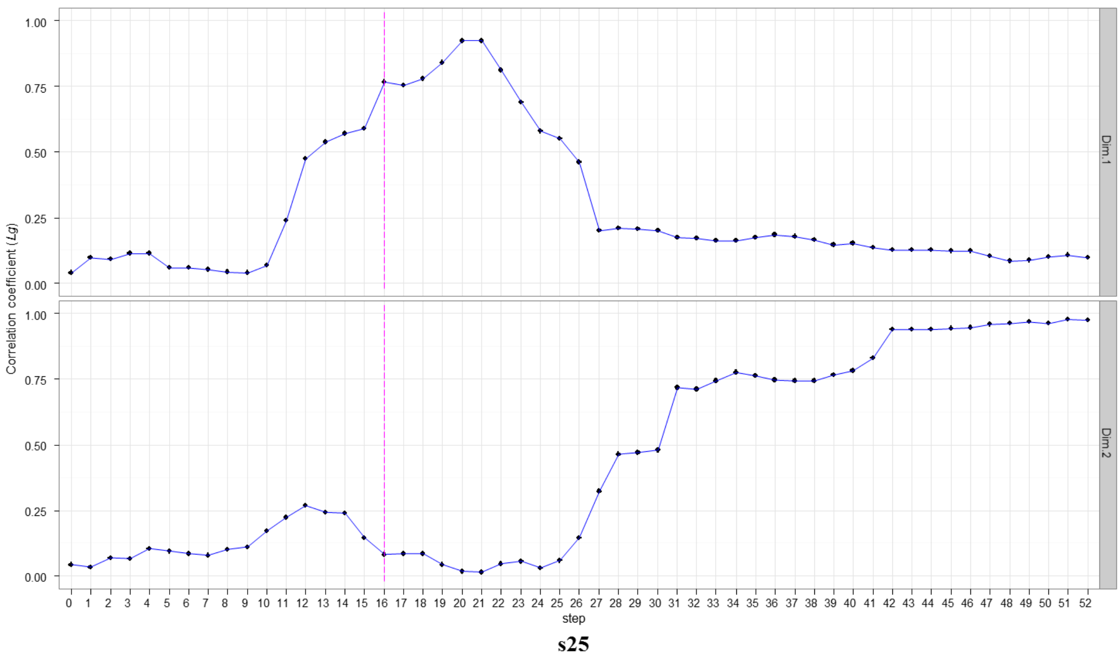


Figure 3.7 Illustration of the learning phase of subject s25 provided by using digit tracking technique

From our point of view, this "aha" moment should be the moment when a subject discovers the whole set of stimuli, rather than a few stimuli. The reason behind of this choice is that: for most of the cases, we cannot know how many dimensions contained in a set of stimuli.

Thanks to digit tracking technique, the "aha" moment can be determined that we denoted *step.learning* in Table 3.1.

Figure 3.7 illustrates the *step.leaning* of subject 25 – a red dash line – along with his/her cognitive process. As we can see, from the moment this subject began the task (step 0) to the "aha" moment (step16)[2], he/she perceived only the first dimension. After discovering all the dimensions of a set of the stimuli, he/she made judgment on the second dimension as the most important dimension to differentiate a set of products. We propose that the period from step 0 to the "aha" moment can be considered as a learning phase – the phase subject gains knowledge from the stimuli.

Learning phase combining with cognitive process might be a useful tool for qualitative observations and/or future investigations relating to subject behaviours on Napping® task (e.g., whether or not subjects make judgments on the most important dimension after assessing all the stimuli).

[2] Technically, the step learning presented in Table 3.1 is taken into account step 0 as the real step (since we want to illustrate step 0 in the representations), hence for subject 25, his/her learning step is 17.

(This page is left blank)

4

Holos: a collaborative environment for holistic methods

From the attempt to its realisation

Au cours des trois premiers chapitres de ce manuscrit, nous avons progressivement développé l'idée selon laquelle l'intégration et l'analyse d'une information supplémentaire telle que le processus cognitif d'un sujet lorsqu'il réalise une tâche de Napping®, sont essentiels à une interprétation fine des résultats. Pratiquement, pour appuyer cette idée, qui constitue la clé de voûte de ce travail de recherche, un environnement de travail *ad hoc*, appelé Holos, a été développé.

Réduit à sa fonction première, Holos peut être simplement vu comme un logiciel qui permet de recueillir des données de Napping® au cours du temps, du début de la tâche à sa toute fin, pour différents types de stimuli, à savoir des mots, des sons, des images, et des vidéos. Au-delà de cette fonction primaire essentielle, Holos est en fait un environnement de travail collaboratif, dans lequel des chercheurs ont la possibilité de partager tout ou une partie de leur expérimentation. C'est cet environnement que nous présentons dans ce chapitre 4, structuré en trois parties complémentaires.

La première partie reprend les éléments principaux du cahier des charges nécessaire au développement de toutes applications. Par analogie, ce cahier des charges s'apparente aux plans conçus par un architecte pour la réalisation d'un édifice. C'est précisément ce rôle que nous avons joué. Le cahier des charges, pensé et rédigé avec minutie, nous a permis d'externaliser le développement de l'environnement Holos, confié à une société de développement informatique. Il fait donc partie intégrante de nos travaux de recherche, puisqu'il est au cœur de la conception de l'environnement Holos.

La seconde partie constitue une mise en situation d'Holos. Cette partie est extraite du manuel d'utilisation d'Holos que les utilisateurs peuvent retrouver sur le site dans lequel se situe l'application.

Enfin, la troisième partie, revêt d'une importance capitale puisqu'elle décrit les fonctions développées au cours de cette thèse, qui permettent le traitement statistique des données issues de l'environnement Holos. Ces fonctions incluent notamment deux fonctions principales. Une première qui correspond au chapitre 2 de ce manuscrit, et qui permet de modéliser et de représenter le processus cognitif d'un sujet à partir de la configuration finale issue de sa tâche de Napping®. Une seconde, qui correspond au chapitre 3 de ce manuscrit, qui permet d'obtenir une représentation du processus cognitif d'un sujet au cours du temps lorsque l'on dispose de données de Napping® temporelles. Ces fonctions ont été intégrées dans le package SensoMineR dédié au traitement de données sensorielles avec R (Lê & Husson, 2008).

Chapter outline

This chapter is an extended version of the following article (see Appendix 4):

Lê, M.T., Brard, M., & Lê, S. (2016, accepted). Holos: A collaborative environment for similarity-based holistic approaches, Behaviour Research Methods.

Abstract

In the present article, we introduce a new collaborative environment, called Holos, in which researchers can carry out experiments based on categorization and Napping® (Pagès, 2005), and share their study resources with the scientific community. Holos is the combination of a software solution for collecting data and a database for storing study resources such as protocols, stimuli and data.

Basically, Holos is an android application with a tactile human machine interface in which subjects can easily conduct experiments using a tablet. Stimuli are displayed as icons that can be dragged with one finger. Stimuli are positioned depending on the way they are perceived. When required, subjects can write down information to describe the stimuli.

As subjects perform the task, the database records their finger movements; more precisely, the trajectories of the icons when they are dragged: this new concept has been called digit-tracking, with reference to eye-tracking. The data recorded enable to study over time what could be assimilated to the cognitive process of a subject during the experiment.

Keywords: Holos; Android application; Digit-tracking; Categorization; Napping®.

4.1 Softwares for Napping® and sorting

4.1.1 An overview

In the last decade, sensory methods in general, sorting and Napping® in particular, have been applied in different research fields. Consequently, new softwares (i.e., tools) appeared on the market in order to take into account that emergent diversity of sensory data. In order to choose the right tool, researchers and practitioners have to think about the way their experiments are usually designed. In particular they have to focus on two criteria, namely the type of stimuli, and the type of locations where subjects conduct the experiment.

Relating to the type of stimuli, we can cite here some examples in which different types of stimuli have been used, such as texts (Bécue-Bertaut & Lê, 2011), images (Courcoux, Qannari, & Faye, 2015), sounds (Berland, Gaillard, Guidetti, & Barone, 2015), or videos (Robitza, Pitrey, & Hlavacs, 2013).

Relating to the type of locations, experiments can take place at sensory central locations, at subjects' home, or even intermediate locations such as supermarkets or shopping centres. Depending on the types of location, different types of technological support will be chosen. For instance, in the case where experiments take place at sensory central locations, an *internal network solution* seems to be the best option. Technically, subjects conduct experiments on computers at the centres, and the data are stored locally. Whereas, in the case experiments take place at subjects' home or at intermediate locations, the technological support is more complicated as it requires a *server-based solution*. Technically, subjects conduct experiments using mobile devices (i.e., a tablet or a smartphone), and the data are stored on an external server using a wireless network (i.e., Wi-Fi).

Table 4.1 presents the current softwares and their supports relating to the type of stimuli and the type of experimental locations: each software supports some (not all) types of stimuli, and three of them use the server-based solution. What could justify such differences?

Originally, TimeSens, Fizz, Logic8, Nappl have been developed for studies in a food context, where stimuli are frequently food products. These stimuli are displayed in these softwares in form of texts or images. Whilst TCL-LabX and NappingPlayer have been developed for studies in psychology and multimedia quality context, where stimuli are sounds and videos. For the technological support, the three softwares that use the server-based solution are TimeSens, Fizz, and Logic8; they are all commercial softwares.

Regarding the notion of temporal data and how to collect them, TCL-LabX is the only one to support this function. However, it appears that TCL-LabX only focuses on the duration of the task and the order of subjects' movements in which the stimuli are assessed, with no particular interest in the notion of cognitive process.

Table 4.1 Current softwares and their supports for studies using Napping® and sorting

N⁰	Software	Stimuli supported				Technology supported	
		Text	Image	Sound	Video	Internal network solution	Server-based solution
1	TimeSens (ChemoSens)	×					×
2	Nappl (Esensorics)	×				×	
3	Fizz (Biosystem)	×	×				×
4	Logic8 (EyeQuestion)	×	×				×
5	TCL-LabX	×		×		×	
6	NappingPlayer			×	×	×	

(×): type of stimuli or technology supported by corresponding software

4.1.2 Holos environment: an overall description

Holos is a new collaborative environment in which researchers can carry out sorting and Napping® experiments and can share study resources. The Holos environment possesses four main functions.

First, within the Holos environment, the experiments can be performed on different types of stimuli, such as texts, images, sounds, and videos. Second, the experiments can take place at subjects' home or intermediate locations. Besides the two must-have functions that may satisfy a wide variety of applications, Holos possesses a third important feature, which allows researchers to share their study resources for Science Exchange. It is expected that by sharing either partially or totally their stimuli and/or their data, researchers can *validate experimental results* via independent replications (i.e., reproducibility initiative), *gather useful information* from different angles (i.e., from other experts), collaborate together (e.g., by conducting cross-cultural studies). Finally, above all, Holos was designed then conceived to collect temporal data issued from Napping®.

Technically, the Holos environment is a server-based data system, which consists of a web-based interface, two applications, and a database. The web-based interface is devoted to researchers in order to design their experiments directly on the server without installing any softwares on their personal computer. The two applications are supported for two main purposes: the first one will facilitate data collection process by which subjects can conduct experiment using Android tablet devices, and the second one allows researchers to revisit all the movements that subjects passed through to complete the task. Finally, the database is located on a server for an easy access of the data wherever the researcher is and as long as he/she is connected. Thanks to the server-based solution, the data collected can be analysed immediately, and study resources can be shared within the scientific community.

In relation to configuration and implementation constraints, the Holos server works on Apache, PHP (version 5.3.3) hosted on Linux, and the database server supports MySQL (version 5). The Android applications work in Jelly Bin (Android version 4.1, 4.2, and 4.3) and KitKat (Android version 4.4). The tablet device used to conduct experiments must to satisfy certain features as its screen resolution of at least 1280 × 800 with pixel density from 160 to 240 ppi.

Researchers, practitioners, and students working in the field of sensory science or other research fields (e.g., psychology, linguistic, product quality, multimedia, etc.) are all the intended users of Holos. Before using Holos, it is worth noting some technical supports for the experiments within the environment. First, Holos supports unlimited space for each user, but the number of stimuli for an experiment is limited to 200 stimuli. Second, for stimuli such as texts, images (formats .png and/or .jpeg), and sounds (formats .mp3 and/or .wav), the size of all the stimuli per experiment is restricted to a maximum of 100MB, with 10MB for each stimulus. Finally, for stimuli such as videos, the streaming technique is used. This technique requires researchers to upload their videos on their personal YouTube channel: the videos will be streamed from the channel to tablet devices when subjects conduct experiments. In relation to study resource security, by default, these resources are stored on the database, and they are not shared within the community. If researchers want to share them, they can share them either partially or totally.

The Holos environment is copyrighted by Tâm Minh Lê and Sébastien Lê from the Unité de Mathématiques Appliquées, Agrocampus Ouest (Rennes, France) and the Sensory Lab, Ho Chi Minh City University of Technology (Ho Chi Minh, Vietnam). It is protected by an end-user license agreement. The agreement grants users a non-exclusive license to use the environment, but requires that the users do not modify, reproduce, license, or sublicense the environment.

4.2 Guidelines for using Holos environment

First, Holos users have to create an account via the website http://napping.agrocampus-ouest.fr. Once the account is validated, they will be able to use the environment.

The main steps for performing an experiment within the Holos environment are similar to those performed in other softwares or environments. First, researchers have to design the experiment (preparing the protocol and the stimuli). Then, researchers recruit subjects to conduct experiment. Finally, they manage the study resources (the stimuli and the data collected). The following sections illustrate how to perform an experiment in the Holos environment.

4.2.1 Design experiment

First, researchers need to login the Holos environment using their valid account via the website http://napping.agrocampus-ouest.fr/login_register.php. Figure 4.1 illustrates four steps to design an experiment, such as: add a new experiment, import the stimuli, write the instruction, and publish the experiment online.

- Add a new experiment: Click button **Add Experiment** (cf. Figure 4.1.a). Certainly, at this first time, there are no experiments created.

- Import the stimuli: Define the type of stimuli as **Image |** Enter a **Name** for the experiment (e.g., Cards with Fractional Factorial Design) **|** Click **Add image | Enter the number of stimuli** (there are 16 cards as stimuli were used) **|** Click **Accept | Upload images** (for each stimulus, we have to upload both a real image with a true size, max. 1020×680 pixels; and a thumb, max. 120×120 pixels) | Click **Submit** to finish the upload images. The result of this step is illustrated in Figure 4.1.b.
- Write the **Instruction** for the current experiment (cf. Figure 4.1.c). This instruction will be appeared on the tablet device when the subjects perform the task.
- Publish the experiment: choose the **Status** as **Publish** | Click **Submit** (cf. Figure 4.1.d). By default, all the experiments are at the **Publish** status, which means: this experiment is ready for all the subjects, who are involved in the subject list (cf. section 4.2.2). In the case that the experiment is under construction and researchers do not allow their subjects to assess, **Status** of the experiment should be set as **Un-publish**.

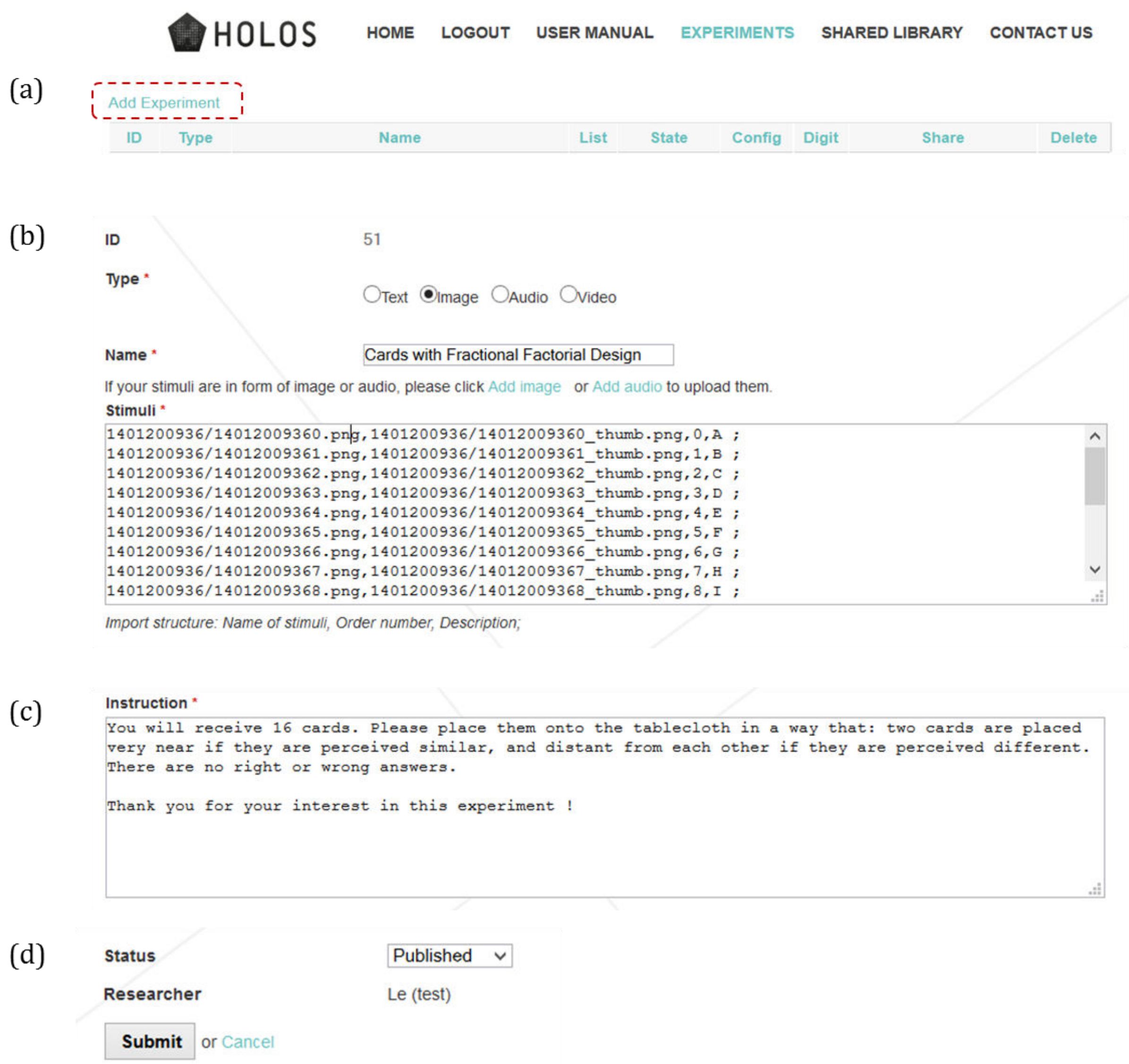


Figure 4.1 Four main steps to design a Napping® experiment on the Holos server

4.2.2 Recruit subjects

For the subjects who will be involved in an experiment, their email addresses need to be filled up in the subject list of corresponding experiment. By default, the subject list is manually filled up by the researchers before experiment time.

Within the Holos environment, email address of one subject and ID of the experiment are considered as the *password* for that subject to assess the experiment. ID number of an experiment is assigned automatically by the Holos server; for instance, ID of the current experiment is assigned as 51 (cf. Figure 4.1.b), it must be provided to subjects before experiment time.

For easily preparing the experiments or blinding of subjects (for studies taken place in intermediate location), email addresses of subjects can be either the real addresses or the virtual ones. However, it is important to note that the email addresses filled up in the subject list need to be the same as those filled up by subjects when conducting experiment (cf. Figure 4.2.b).

In order to create a subject list, researchers: Click on the **Subject List |** Add the email addresses of subjects in the **Email tab**. In the case where the real email addresses are used, researchers can send the invitation and information of experiment via subjects' emails by clicking **Send email**.

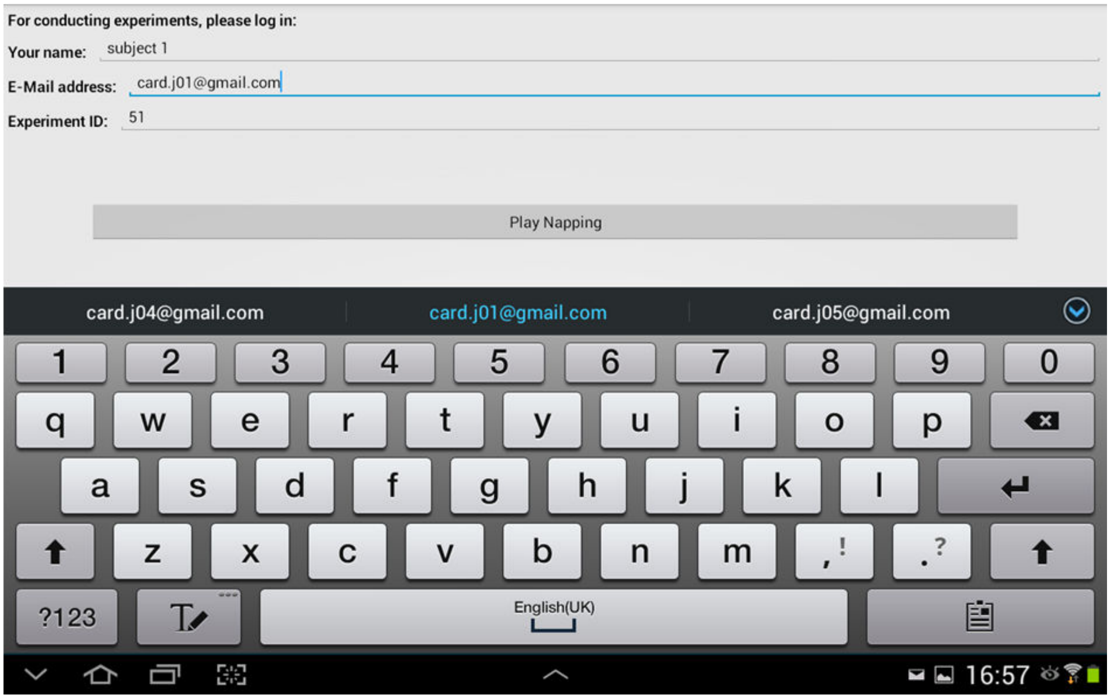


Figure 4.2 An initial assessment of the experiment by a subject using tablet device

4.2.3 Ask subjects to conduct experiment

This step is devoted to the subjects who will conduct Napping® experiments on the tablet device. Firstly, researchers ask subjects to install an Android application, called NSubject (i.e., **N**apping for **Subject**), which is available at the website http://napping.agrocampus-ouest.fr. Its installation is similar to that of other applications in Google Play store. Once the NSubject application is installed, subjects can assess the experiment as follows: Click **NSubject icon** on the tablet | Enter

subject's information including: Name, Email address, and ID of the experiment, then click **Play Napping** (cf. Figure 4.2).

It is worth noting that at this first step the stimuli are *randomly* presented to each subject. To conduct the experiment, subjects can use five hand features supported by the NSubject application, such as: *drag* to move, *pinch* to zoom in, *spread* to zoom out, *tap* on the icon to select or group, and *double-tap* to deselect or ungroup the images. Once the task is completed, subjects can send the results by clicking **Finish** button. From this step, the data is stored on the Holos server.

4.2.4 Manage study resources

Within the Holos environment, study resources (including: the stimuli, temporal data, as well as final configuration of the subjects) are stored on the server. Researchers have a direct access to their data in order to analyse them, they can also share their study resources. In addition to this, Holos support a visual tool, named NAdmin (i.e., **N**apping for **Admin**istrator), it is another Android application allowing to revisit subjects' movements during the task.

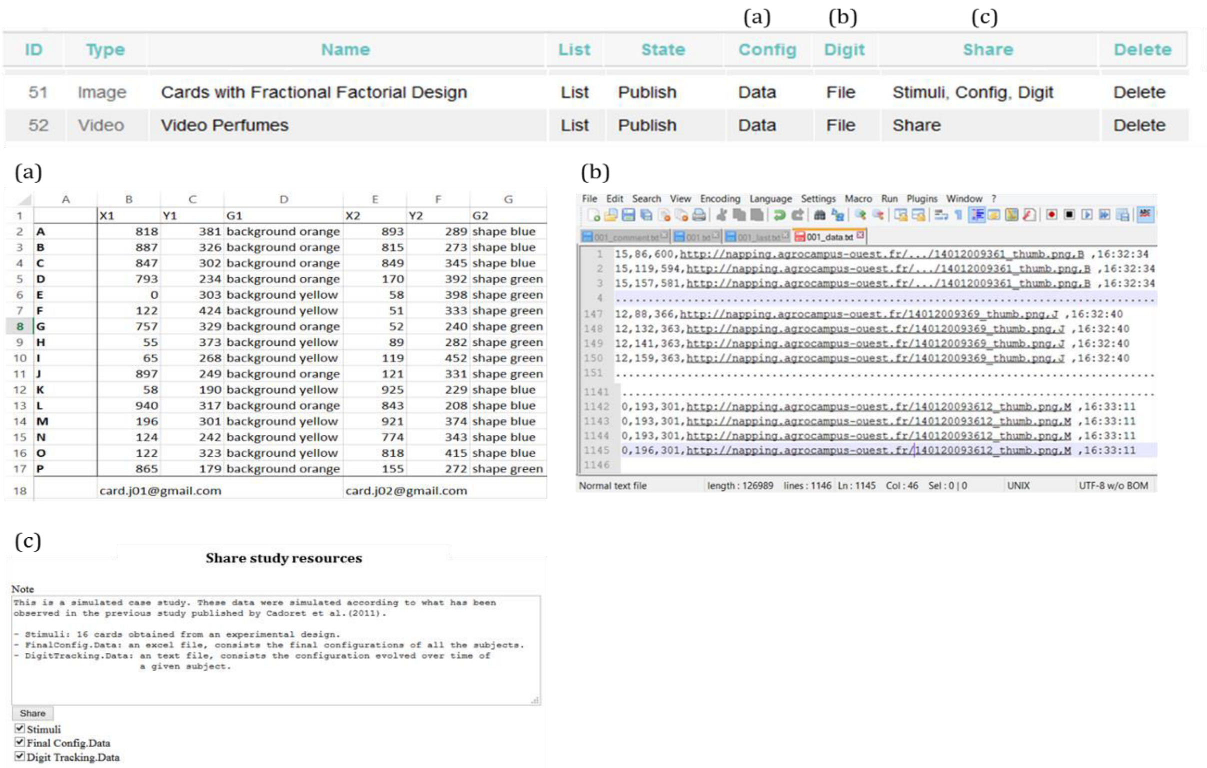


Figure 4.3 Study resources stored on the Holos server: the final configuration of all the subjects (a), the temporal data of each individual subject (b), and together with the stimuli, they can be shared within the scientific community (c)

In terms of data collected, the final configurations can be assessed by clicking on the tab **Config | Data**. It is an excel file, which consists of the coordinates of the stimuli at the final step obtained by all the subjects (cf. Figure 4.3.a). The temporal data can be assessed by clicking on the tab **Digit | File**. This folder consists of many text files, each corresponds the coordinates recorded over time of each individual subject (cf. Figure 4.3.b). In order to share study resources within the scientific community, researchers: click on the tab **Share**, then choose study resources (stimuli, final configurations, and/or the temporal data) to share (cf. Figure 4.3.c).

Beyond the Holos server, temporal data issued by individual subject is also stored on the tablet device. To revisit the movements of a given subject, researchers need to install NAdmin application, available at the website http://napping.agrocampus-ouest.fr. Then: activate the **NAdmin |** choose the **Subject** | click **Start view**.

4.3 Analysing (temporal) data issued from Napping® using SensoMineR

This section presents two couples of functions for analysing (temporal) data issued from Napping® methods. In this:

- Functions `cpdm()` and `plotcpdm()` aim to obtain the graphical outputs of cognitive process based on divisive model (cf. Chapter 2).
- Functions `traject()` and `dimtracking()` aim to obtain the graphical outputs of cognitive process evolving over time (cf. Chapter 3).

These functions are integrated in the R package SensoMineR. In this section, they will be illustrated through a simulation data on cards. This data was chosen for two reasons: first, it is a pedagogical example; and second, it is linked with the illustration presented in the previous section.

4.3.1 Cards data

Cards data was simulated according to what Cadoret et al. (2011) observed when children conducted descendant hierarchical sorting task on 16 cards.

For the simulation process, we considered 5 types of subjects. Type S1 (40 subjects) who divided stimuli into 2 groups based on the background color. Type S2 (30 subjects) who divided stimuli into 2 groups based on the shape color. Type S3 (5 subjects) who divided stimuli into 4 groups based on both background color and shape color. Type S4 (5 subjects) who divided stimuli into 2 groups based on the contour. Finally, Type S5 (30 subjects) who divided stimuli into 2 groups based on: first the background color, then the shape color (cf. Figure 4.4).

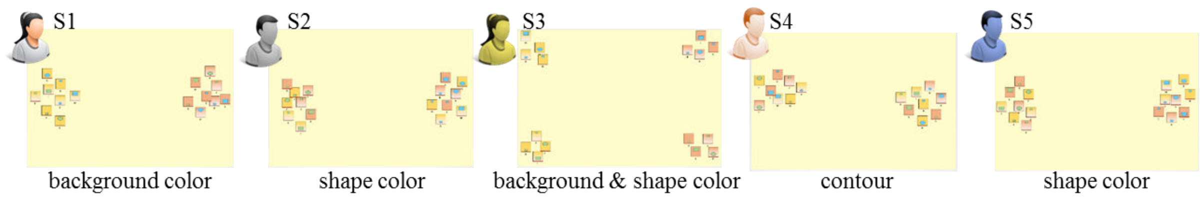


Figure 4.4 Five types of subjects who perceived stimuli in different ways

This simulation data was separated into two datasets:

- `data(cards.fico)`: a data frame, which consists of the final configurations of all the subjects (110 subjects). This data will be used to illustrate the functions `cpdm()` and `plotcpdm()`.
- `data(cards.tempo)`: a data frame, which consists of the temporal data of all the subjects. This data will be used to illustrate the functions `traject()` and `dimtracking()`.

4.3.2 Functions for obtaining cognitive process using divisive model

4.3.2.1 *Function* `cpdm()`

This function will perform the divisive analysis on the Napping® data of each individual subject (i.e., the *X*- and *Y*- coordinates) with the number of successive steps is pre-defined. The structure of `cpdm()` is given below:

```
cpdm(X, nb.step)
```

Arguments:

- `X`: a data frame, which consists of the final configurations of all the subjects.
- `nb.step`: the number of successive steps (i.e., levels) will be analysed.

Output of this function is a list of data frames, including:

- `Napping.data`: a data frame, initial Napping® data (quantitative variables).
- `Steps.of.all.subjects`: a data frame, which is the results obtained by divisive model (qualitative variables).
- `CPDM.data`: the combination of two data frames, `Steps.of.all.subjects` and `Napping.data` (cf. Figure 4.5). This data will be used to depict cognitive process of the subjects using function `plot.cpdm()`.

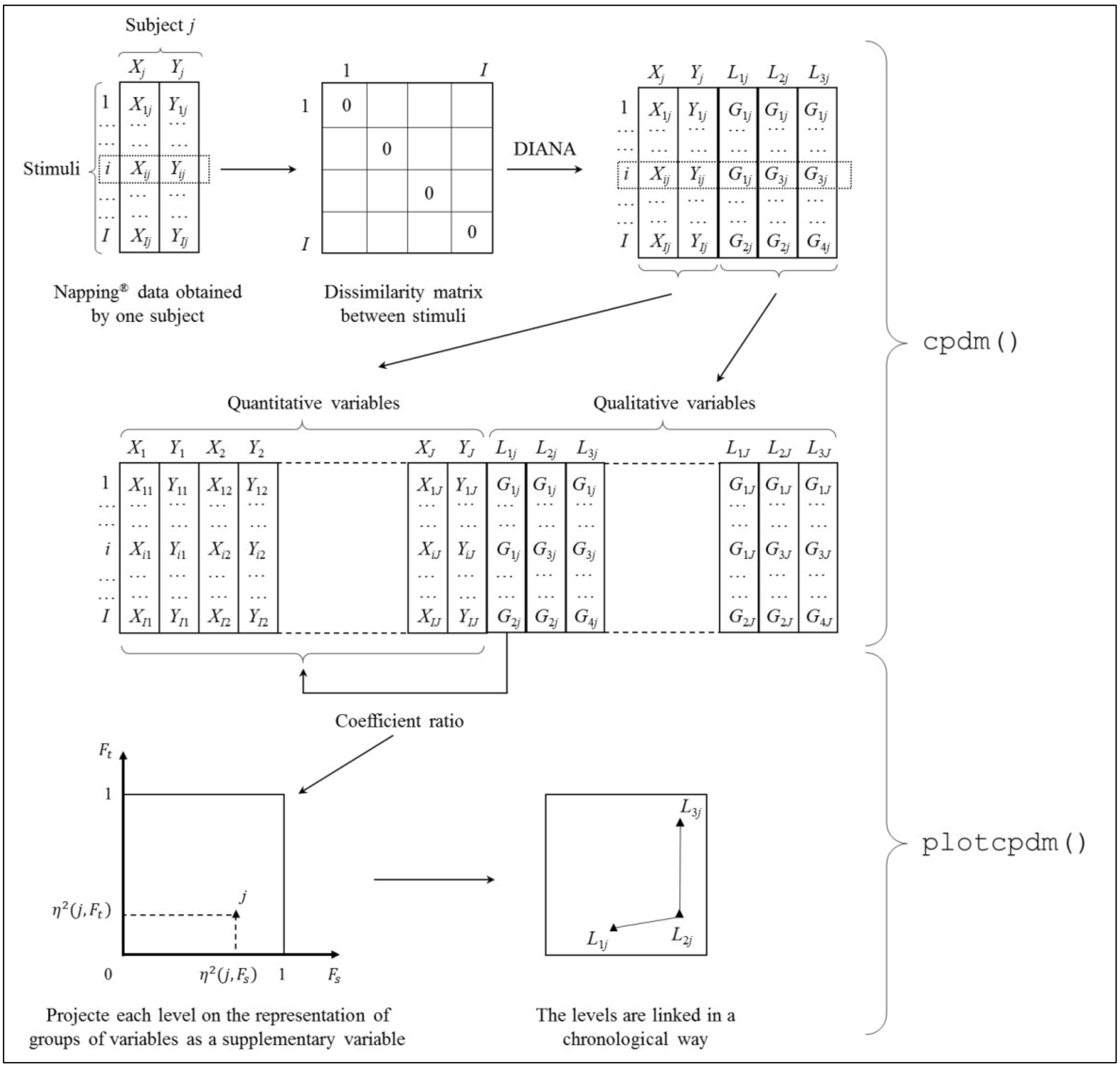


Figure 4.5 Statistical process and the outputs of functions `cpdm()` and `plotcpdm()`

4.3.2.2 Function `plotcpdm()`

In order to depict cognitive process of each individual subject, `plotcpdm()` performs three steps on `CPDM.data` (cf. Figure 4.5). First, correlation ratios are calculated between each level of the hierarchy and two principal dimensions provided by MFA (as presented in Chapter 2, two principal dimensions were provided when performing MFA on initial Napping® data). Second, all levels of hierarchy of each subjects are then projected on the representation of the groups of variables. Finally, they are linked in a chronological order. The structure of `plotcpdm()` is given below:

```
plotcpdm(res.cpdm, subject, axes = c(1, 2))
```

Arguments:

- `res.cpdm`: results obtained by `cpdm()`.
- `subject`: individual subject whose cognitive process will be depicted.

- `axes`: principal dimensions provided by MFA, in which cognitive process will be depicted, by default Dim 1 & 2 are chosen.

Output

- A graph, which depicts a cognitive process of an individual subject.

Examples

```
> # Require package ---------------------------------------------------
> library(FactoMineR)
> library(SensoMineR)
> library(cluster)
> library(grDevices)
> # Load data ---------------------------------------------------------
> data(cards.fico)     # final configurations of subjects
> # Execute function cpdm() and plotcpdm()---------------------------
> cards.cpdm <- cpdm(cards.fico, nb.step=3)
> plotcpdm(cards.cpdm, subject=1)  # s1: 40 subjects
> ## Plot: subject=41 (type s2, 30 subjects), subject=71 (type s3, 5
        subjects), subject=76 (type s4, 5 subjects), subject=81
        (type s5, 30 subjects)
```

Results

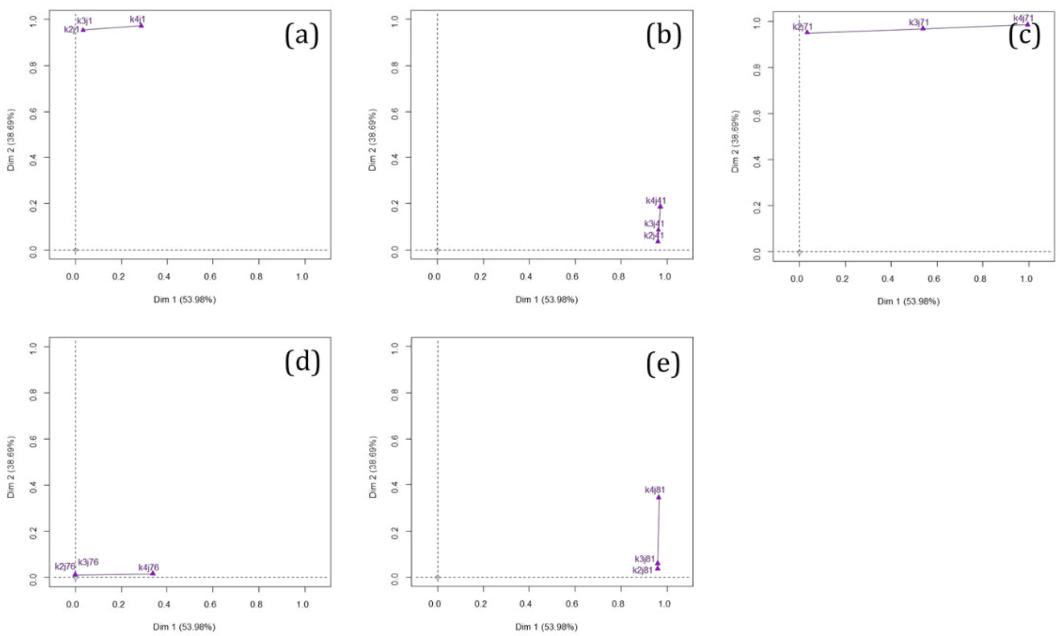


Figure 4.6 Cognitive process of the five types of subjects depicted using function `cpdm()`

4.3.3 *Functions for obtaining cognitive process evolving over time*

4.3.3.1 *Function* `traject()`

This function performs three main steps on temporal data issued from Napping®. First, an MFA is performed on the final configuration of all the subjects to obtain the reference framework. Then, the matrices distance between stimuli evolving over time, denoted $W_j(t)$, are projected as supplementary elements on the representation of the groups of variables by means of Lg measure (cf. section 3.2.2). Finally, the presentations of $W_j(t)$ are linked in a chronological order.

The structure of `traject()` is given below:

```
traject(data, juge, axes, plot.graph=TRUE)
```

Arguments

- `data`: temporal data issued from Napping®, e.g. `cards.tempo`.
- `juge`: an individual whose cognitive process evolving over time will be depicted.
- `axes`: the dimensions which constitute the representation of the groups of variables provided from MFA.
- `plot.graph`: boolean, if TRUE (or by default) the graph will be displayed.

Output

- A graph, which depicts a cognitive process evolving over time of individual subject.
- A data frame, which consists of Lg measures between $W_j(t)$ and the components (F_s, $s \geq 1$) provided by MFA.

Example

```
> library(FactoMineR)
> library(SensoMineR)
> library(ggplot2)
> library(reshape2)
> library(stringr)
> library(car)
> library(cluster)
> library(grDevices)
> # Load data -------------------------------------------------------
> data(cards.tempo)
> cards.tempo[1:32,]   # structure of temporal data
  subject step stimulus     time coordX coordY
1       1    0        M 16:32:28      0      0
2       1    0        K 16:32:28      0     80
3       1    0        P 16:32:28      0    160
4       1    0        D 16:32:28      0    240
```

```
5         1   0       H 16:32:28     0    320
6         1   0       A 16:32:28     0    400
7         1   0       G 16:32:28     0    480
8         1   0       O 16:32:28     0    560
9         1   0       L 16:32:28    80      0
10        1   0       N 16:32:28    80     80
11        1   0       I 16:32:28    80    160
12        1   0       E 16:32:28    80    240
13        1   0       J 16:32:28    80    320
14        1   0       F 16:32:28    80    400
15        1   0       C 16:32:28    80    480
16        1   0       B 16:32:28    80    560
17        1   1       B 16:32:35   887    326
18        1   2       C 16:32:38   847    302
19        1   3       J 16:32:42   897    249
20        1   4       L 16:32:45   940    317
21        1   5       A 16:32:47   818    381
22        1   6       D 16:32:50   793    234
23        1   7       G 16:32:52   757    329
24        1   8       P 16:32:55   865    179
25        1   9       O 16:33:00   122    323
26        1  10       F 16:33:01   122    424
27        1  11       H 16:33:02    55    373
28        1  12       E 16:33:03     0    303
29        1  13       I 16:33:04    65    268
30        1  14       N 16:33:06   124    242
31        1  15       K 16:33:09    58    190
32        1  16       M 16:33:11   196    301
> # Function traject() ---------------------------------------------
> corr.j1  <- traject(cards.tempo, juge=1, axes=c(1,2))
```

Results

```
> corr.j1[,1:3]  # Lg measure values on 3 principal components
        Dim.1  Dim.2  Dim.3
Step0  0.0029 0.0618 0.0057
Step1  0.0834 0.0663 0.0449
Step2  0.1587 0.0891 0.0017
Step3  0.0533 0.1998 0.0162
Step4  0.1290 0.2859 0.0696
Step5  0.2450 0.3446 0.0110
Step6  0.1506 0.4919 0.0013
Step7  0.0806 0.6530 0.0380
Step8  0.0357 0.9509 0.0073
Step9  0.0428 0.9426 0.0030
```

```
Step10 0.0406 0.9447 0.0026
Step11 0.0383 0.9489 0.0027
Step12 0.0442 0.9480 0.0018
Step13 0.0483 0.9480 0.0009
Step14 0.0485 0.9436 0.0008
Step15 0.0515 0.9387 0.0017
Step16 0.0660 0.9336 0.0012
```

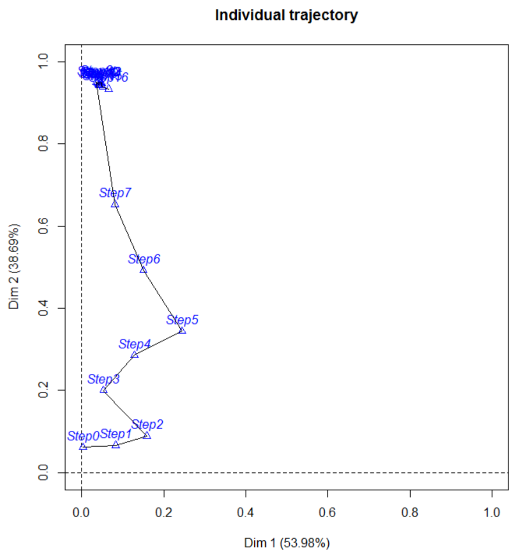


Figure 4.7 Cognitive process evolving overtime of subject 1 depicted by function `traject()`

4.3.3.2 Function `dimtracking()`

In order to track the cognitive process perceived by subjects in each dimension, function `dimtracking()` should be used. Its structure is given below:

```
dimtracking(corr.data, num.dim, smooth=FALSE, span=span)
```

Arguments

- `corr.data`: results obtained by function `traject()`.
- `num.dim`: a vector number determining the numbers of dimensions that will be tracked.
- `smooth`: boolean, if TRUE, adding a smoother to a plot. Here, the method *loess* is used to fit a smoother to the small data set where the number of steps performed by a subject is rarely greater than 1,000.
- `span`: if smooth=TRUE, the span parameter must be indicated to control the wiggliness of the line. For more information, see package ggplot2 (Wickham, 2009).

Output

- A graph, which depicts cognitive process of an individual subject tracked on each dimension.

Example

```
> # Function dimtracking() ------------------------------------------
> dimtracking(corr.j1, num.dim=1:3)
>
> # dimtracking() with learning phase
> res.sum <- suminfo(cards.tempo)
> j1.learning <- res.sum$Step.Time$Step.learning[1]
> corr.j1  <- traject(cards.tempo,juge=1, axes=c(1,2),
        plot.graph=FALSE)
> p <- dimtracking(corr.j1, num.dim=1:3)
> p + geom_vline(xintercept = j1.learning, colour="magenta",
        linetype="longdash")
```

Results

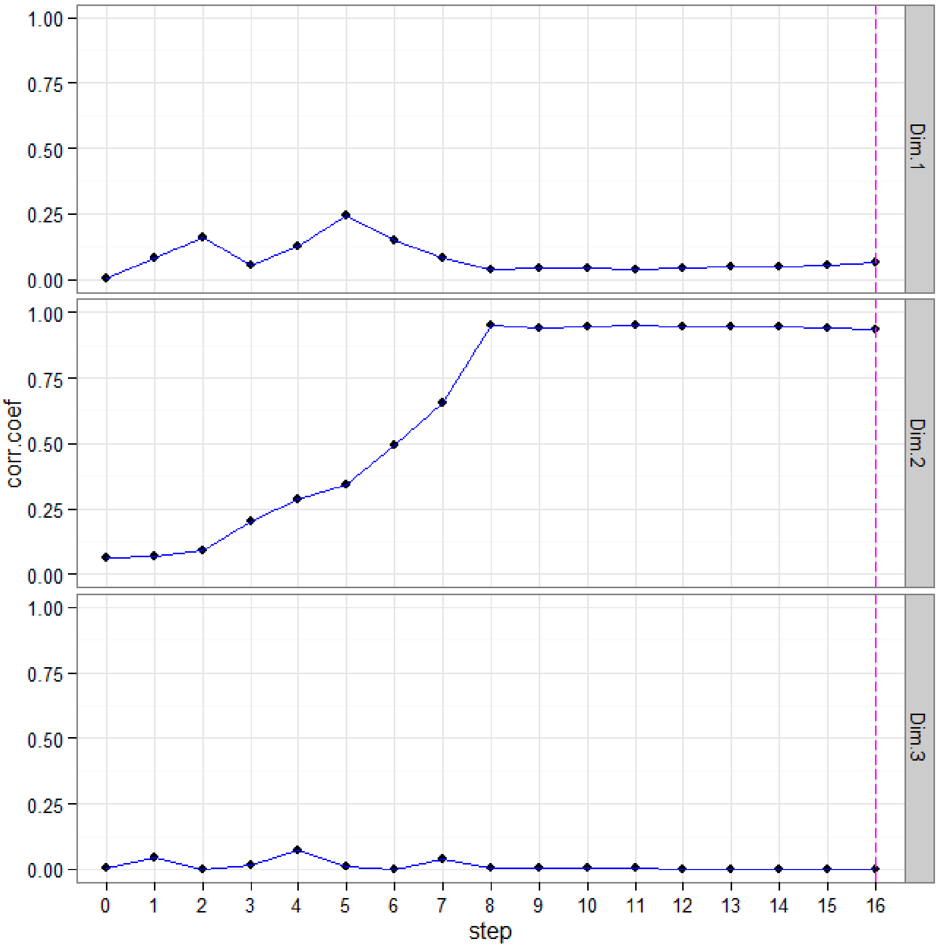


Figure 4.8 Cognitive process of subject 1 tracked on the three first dimensions depicted by function `dimtracking()`

Discussion

La difficulté d'une thèse sur articles réside paradoxalement dans sa rédaction. Les articles et donc les chapitres sont certes écrits, mais il reste ensuite à les articuler et à trouver le liant qui permet de traduire parfaitement la pensée de son auteur. Cette difficulté s'explique en partie, du fait que le document est un assemblage d'articles pouvant se lire de façon presque indépendante. Dans notre cas, nous avons voulu, tout au long de ces trois ans, développer une méthodologie et l'appliquer afin de la comprendre de façon pratique et de la valider. Cette compréhension a été riche car les stimuli sur lesquels nous avons travaillé ont été très différents d'une expérimentation à l'autre. C'est précisément ce qui peut troubler le lecteur de cette thèse qui finalement peut se perdre à travers cette richesse d'exemples et en oublier notre proposition méthodologique.

La discussion que nous menons dans cette partie est l'occasion de reprendre les éléments méthodologiques que nous proposons dans cette thèse, de façon complétement indépendante des exemples. Cette discussion va s'articuler en cinq points. Un premier point concernera le positionnement du Napping® par rapport à d'autres méthodes de recueil de données sensorielles. Ceci nous amènera, dans un second point, à réfléchir sur la donnée elle-même : que peut-on espérer trouver derrière une expérience de type Napping®. Dans un troisième point, nous verrons comment exploiter l'information supplémentaire, supposée cachée, dont nous avons parlé dans le point deux. Le quatrième point est l'occasion de discuter de cette information cachée et que nous avons finalement réussi à découvrir grâce à l'application Holos que nous avons développée au cours de notre thèse. Enfin le cinquième point, nous permettra de nous projeter dans ce qu'on peut finalement attendre de notre proposition méthodologique.

Premier point – Positionnement du Napping® par rapport à d'autres méthodes de recueil de données sensorielles

Dans l'article originel de Jérôme Pagès paru en 2005, le recueil de données sensorielles du Napping® est présenté au sujet en deux temps. Dans un premier temps, le principe de la tâche est expliqué au sujet, puis dans un second temps, on lui présente le protocole expérimental.

Le principe consiste en un recueil de données de distances inter-stimuli, comme l'indique d'ailleurs le titre de l'article : *« Collection and analysis of perceived product inter-distances using multiple factor analysis : Application to the study of 10 white wines from the Loire Valley »*. Chaque sujet compare les stimuli selon ses propres critères puis se fait une idée des distances inter-stimuli, tous critères confondus. Il s'agit ici d'une évaluation globale des produits. C'est en ce sens que la donnée issue de la tâche de Napping® peut être considérée comme holistique. Chaque stimulus est considéré en tant que tel, c'est-à-dire comme un tout et non comme une somme de composantes sensorielles. Cette caractéristique n'est pas propre au Napping®. En effet, on peut également considérer que la comparaison par paires ou le tri fournissent des données de nature holistique. De manière générale, on peut considérer comme holistique tous recueils sensoriels basés sur la notion de distances inter-stimuli.

Le protocole est *spécifique* au Napping®. Il est demandé au sujet de positionner les stimuli selon leur proximité perçue, telle que deux stimuli sont proches s'ils ont été perçus comme proches sensoriellement ; deux stimuli sont éloignés s'ils ont été perçus comme différents sensoriellement. De plus, il est *explicitement* demandé au sujet d'utiliser tout l'espace dont il dispose pour exprimer les distances sensorielles qu'il perçoit à travers les stimuli. De façon schématique, le protocole du Napping® pour un sujet peut être comparé à l'algorithme de l'Analyse en Composantes Principales (ACP). L'objectif de l'ACP est de visualiser la diversité d'un ensemble de points sur un plan de projection qui maximise l'inertie du nuage de points projetés (Husson, Lê, & Pagès, 2011). On peut traduire la première partie du protocole du Napping® comme une projection d'une matrice de distances inter-stimuli sur un plan à 2 dimensions. De même, on peut traduire la seconde partie du protocole comme une contrainte sur cette projection, à savoir une contrainte de maximisation de la matrice de distances inter-stimuli positionnés sur le plan. Par conséquent, le Napping® peut être comparé à l'ACP dans le sens où chaque sujet, à travers une matrice de distances inter-stimuli projetées sur la nappe, révèle ses principales dimensions sensorielles de variabilité. Par construction, la première dimension est celle pour laquelle la diversité entre les produits est la plus importante. On retrouve ici la capacité à *révéler* et à *ordonner* du Napping®.

L'analogie que nous venons de faire avec l'ACP pour présenter la tâche de Napping® va nous permettre de la positionner vis-à-vis du QDA® d'une part, et du Projective Mapping d'autre part, afin de mieux la comprendre.

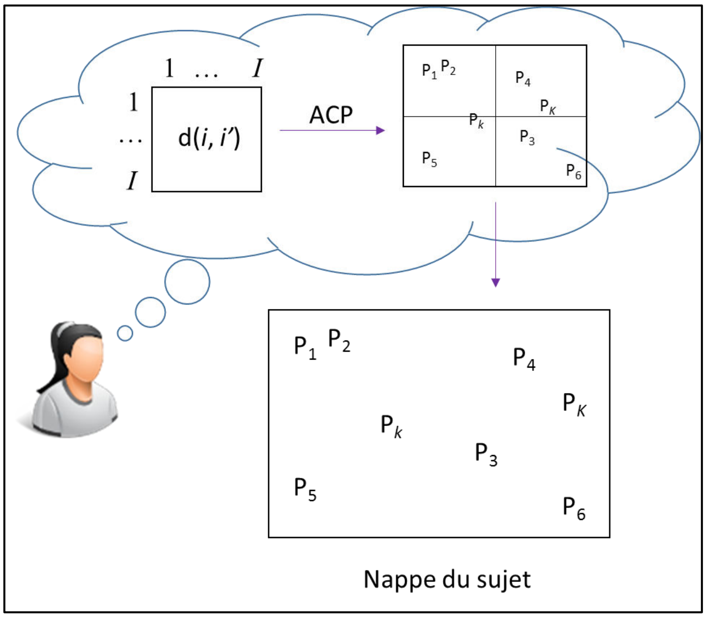


Figure D-1 De façon schématique, le protocole du Napping est considéré comme l'algorithme de l'Analyse en Composantes Principales où le sujet projete d'une matrice de distances inter-stimuli sur un plan à deux dimensions

La différence fondamentale entre le Napping® et le QDA® réside dans leur point de départ. En Napping®, le sujet dispose d'un ensemble de stimuli, alors en QDA®, il dispose d'un ensemble de stimuli et d'*une liste de descripteurs*. En Napping®, le point de vue d'un sujet est non-fixé *à priori*, alors qu'il l'est en QDA®. Par conséquent, le Napping® peut générer une multiplicité de points de vue, contre un seul en QDA®.

Ces points de départ conditionnent l'usage que l'on a de ces deux méthodes. En Napping®, à travers les oppositions qu'un sujet génère entre des produits, il révèle des dimensions sensorielles qui pour lui sont importantes. En QDA®, le sujet mesure des intensités sur des descripteurs pour lesquels il a été entraîné. En pratique le Napping® sert donc à *révéler* un ensemble d'images sensorielles pour un ensemble de stimuli donnés, alors que le QDA® permet d'*obtenir* l'image la plus précise possible pour un ensemble de stimuli donnés et pour un point de vue donné. La difficulté du Napping® réside dans la richesse des données qu'il fournit, puisqu'il s'agira de faire une synthèse de cette diversité de points de vue et de l'interpréter.

Depuis l'article originel de Pagès (2005), de nombreux articles ont été écrits sur le Napping® et sur son utilisation. On a vu réapparaître le concept de Projective Mapping amené quelques années auparavant par Risvik et al. (1994). Petit à petit, on a assisté à une dérive sémantique du terme Napping® utilisé dans un premier temps en tant que tel, puis conjointement avec le Projective Mapping. Enfin, presque définitivement remplacé par le Projective Mapping. La différence fondamentale entre le Napping® et le Projective Mapping réside également dans leur point de départ.

Dès les premières lignes de leurs articles, alors que Pagès (2005) oppose les méthodes où l'on mesure les distances inter-stimuli aux méthodes de profil classique ; Risvik *et al.* (1994) positionne le Projective Mapping au même niveau que le QDA® et la comparaison par pairs. La raison de ce positionnement est due aux différences entre les objectifs.

Pour Pagès, il s'agit de trouver une alternative à des méthodes de recueil inter-distance, dans le but, d'*ordonner* des dimensions sensorielles perçues comme importantes par le sujet. Pour Risvik, il s'agit de trouver une méthode alternative au QDA® et à la comparaison par paires, dans le but, d'*obtenir* un espace produit. De cette différence de positionnement et d'objectif découlent des différences au niveau des protocoles et des analyses statistiques. Trois points nous paraissent importants :

- La dimension de la nappe et notamment le rapport de 1,5 entre la largeur et la hauteur, spécifié par Jérôme Pagès.
- L'utilisation de tout l'espace disponible lors de l'évaluation des distances inter-stimuli. Ce point figure dans le protocole proposé par Pagès, mais il ne figure pas dans celui proposé par Risvik.
- Enfin, le traitement par AFM, proposé par Pagès, en considérant des groupes de coordonnées non-normés, qui prend bien en compte la spécificité des données recueillies, ce que ne fait pas Risvik.

Chacun de ces trois points contribuent à *ordonner* les facteurs révélés par le sujet.

À travers la réflexion du recueil du Napping®, s'il fallait retenir deux idées importantes de ce recueil, nous choisirions celles-ci :

- Tout d'abord, le Napping® permet de *révéler* les principales dimensions sensorielles de variabilité d'un sujet pour un espace produit donné.

- Ensuite, le Napping® permet d'*ordonner* les principales dimensions sensorielles de variabilité d'un sujet pour ce même espace produit.

Ces 2 idées sont indissociables car c'est l'association des 2 qui a le plus de valeur en pratique. En effet : Quel est l'intérêt des dimensions sensorielles si on ne peut pas les ordonner selon leur importance relative ? Même si les dimensions sensorielles sont ordonnées, quel est leur intérêt si on ne sait pas clairement les identifier ?

Deuxième point – Que peut-on espérer trouver derrière des données de Napping®

Figure D-2 illustre un exemple simple où le Napping® a permis de révéler que les dimensions sensorielles exprimées par ce sujet sont la forme et la couleur. Il a également permis de déterminer que le sujet a accordé plus d'importance à la forme qu'à la couleur. Cependant, ces deux caractéristiques peuvent être mieux interprétées si elles sont associées à des informations supplémentaires.

Le seul fait de révéler ne suffit pas à interpréter les dimensions du Napping®. Pour révéler correctement les dimensions sensorielles, il faut avoir soit une connaissance experte sur les produits soit, une description des produits par le sujet ce qui n'est pas obligatoire. Cette description permet d'interpréter le point de vue du sujet, et renforce l'information qu'il fournit à travers les dimensions. C'est par exemple le cas des données de verbalisation qui, une fois collectées, permettent de mieux comprendre les dimensions sensorielles révélées par le Napping®.

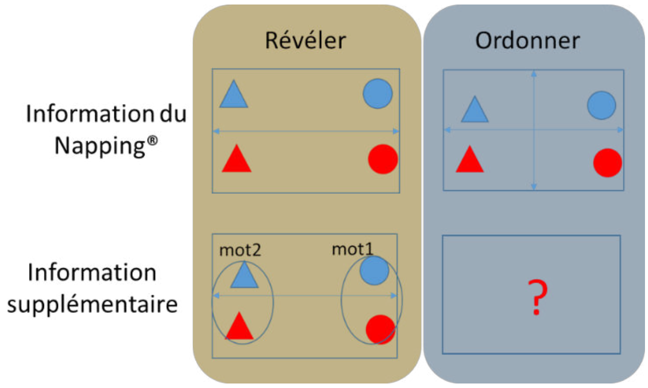


Figure D-2 Réflexion d'information ajoutée pour mieux déterminer l'importance relative de dimensions sensorielles fournies par un sujet

De la même façon, il peut être justifié d'apporter une information supplémentaire pour mieux déterminer l'importance relative de ces dimensions sensorielles. Qu'en est-il de la capacité à ordonner les dimensions sensorielles ? Peut-on renforcer la croyance que l'on a dans l'importance relative d'une dimension par rapport à une autre ? Quelle information supplémentaire faudrait-il recueillir pour mieux appréhender l'importance d'une dimension par rapport à une autre ?

Ces questions se posent naturellement alors même que les dimensions de la nappe devraient nous garantir, en théorie, une première dimension 1,5 fois plus importante que la deuxième.

L'illustration présentée dans la Figure D-3 démontre le fait que la configuration finale des sujets n'est pas suffisante pour interpréter l'importance relative des dimensions sensorielles.

À première vue, le sujet 1 et le sujet 2 ont mis en évidence les mêmes dimensions sensorielles et les ont ordonnées de la même façon : c'est-à-dire la forme sur la largeur et la couleur sur la hauteur. Néanmoins, on remarque que le sujet 1 a fourni cette configuration finale de façon directe. Le sujet 2, lui, a d'abord séparé les stimuli selon leur couleur puis selon leur forme. Il a ensuite changé d'avis pour finalement utiliser la même configuration finale que le sujet 2. Que dire de l'opposition forme puis couleur perçue par ces deux sujets qui aboutissent à une même configuration finale ? Le changement d'avis au cours du temps pour le sujet 2 doit-il renforcer la dimension couleur par rapport à celle de la forme ou, au contraire, doit-il renforcer la dimension forme par rapport à celle de la couleur ?

Partant de ce constat, quelle information supplémentaire peut être utilisée pour mieux appréhender la capacité à ordonner les dimensions sensorielles d'un sujet ? La réponse à cette question se trouve dans l'exemple précédent, où la seule façon de distinguer les nappes fournies par les deux sujets 1 et 2 est de prendre en compte la suite des étapes qui ont menées ces deux sujets à leur configuration finale. L'ensemble des étapes conduisant à la configuration finale sera, par la suite, assimilé au processus cognitif du sujet lors de la tâche de Napping®. Rappelons qu'un processus cognitif peut être défini de la façon suivante : *« ensemble des différents modes à travers lesquels un système traite l'information en y répondant par une action »* (Wikipedia, 2015 ; APA Glossary, 2015).

La partie suivante présente comment recueillir et traiter du comportement (ou processus cognitif) du sujet pour mieux comprendre des données de Napping®.

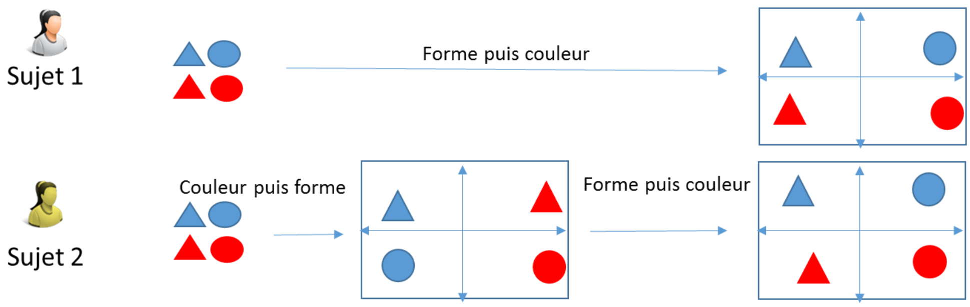


Figure D-3 Illustration des deux sujets qui peuvent fournir la même configuration finale sans utiliser le même processus cognitif

Troisième point – Comment exploiter l'information inhérente aux données de Napping® relatives au comportement du sujet

Il apparait donc que la suite des étapes ayant mené un sujet à sa configuration finale peut constituer une information supplémentaire pour mieux comprendre les données de Napping®. Ces données peuvent être recueillies en direct, c'est-à-dire au moment où le sujet réalise sa tâche ou en différé, c'est-à-dire après que le sujet ait réalisé sa tâche.

La deuxième option, bien qu'elle n'ait pas été testée n'a pas été retenue. Demander à un sujet de se rappeler des étapes qui l'ont amené à sa configuration finale ne nous semblait pas réalisable, que ce soit pour des raisons de fatigue sensorielle ou pour des problèmes de mémoire. C'est donc la première option qui a été retenue, et qui a nécessité le développement d'un logiciel adapté.

Les fonctionnalités de ce logiciel ont été déterminées en fonctions des contraintes expérimentales suivantes. Nous voulions des données :

1. Recueillies en direct pour suivre la suite des étapes du sujet pendant la tâche
2. Pouvant provenir de diverses natures de stimuli (par exemple, des vidéos, des sons, des photos, et des concepts) pour pouvoir étendre le Napping® à plusieurs domaines
3. Pouvant être recueillies à distance

Pour répondre à ces contraintes, il a été décidé de développer un logiciel alliant un recueil de données de Napping® sur tablette et un enregistrement des données sur un serveur. Le développement d'une application étant un processus extrêmement couteux en temps il nous a semblé important de travailler avec des données simulées dans un premier temps, en attendant de recueillir de vraies données associées au processus cognitif des sujets. Ces données simulées avaient pour objectif de servir de support de réflexion lors du développement de l'application et des solutions d'analyse statistique. Comment simuler ces données test ?

À l'image de la statistique Bayésienne qui utilise de l'information *à priori*, pour simuler les données associées au comportement du sujet nous nous sommes basés sur l'hypothèse comportementale suivante : *« face à un grand nombre de stimuli un sujet adopte une stratégie descendante afin d'évaluer des distances inter-stimuli »*. Cette hypothèse est extrêmement restrictive. Ce type de comportement ne peut correspondre qu'à un sous-ensemble réduit de comportements possibles lors d'une tâche de Napping®.

Cependant dans un contexte où le nombre de stimuli est relativement élevé, c'est ce que nous avons observé pour une grande majorité des sujets interrogés. Il semblerait que dans un tel contexte cette hypothèse soit réaliste. La contrainte dictée par le nombre de stimuli impose au sujet de ne pas les évaluer 2 à 2 du fait du trop grand nombre de distance inter-stimuli à mémoriser, autrement dit du fait de la quantité d'informations à retenir. Il semblerait que le sujet appréhende les stimuli dans leur ensemble comme un tout homogène. Il fait alors ressortir les traits saillants de l'espace des stimuli en les divisant de façon séquentielle en sous familles hiérarchisées. La simulation des données sous la contrainte de cette hypothèse comportementale nous a amenés à utiliser un algorithme divisif (DIANA) sur la matrice de distances obtenue grâce au Napping®. Cette méthode consiste à construire un arbre hiérarchique descendant à partir de la matrice de dissimilarité entre des stimuli obtenue avec le Napping® (cf. Figure D-4).

La coupure de l'arbre hiérarchique est déterminée par le nombre de groupes constitués sur la nappe. Comment déterminer le niveau de coupure ? Afin de déterminer le nombre de groupes, nous sommes partis de l'hypothèse suivante : « chaque dimension de la nappe ne peut représenter qu'une opposition entre 2 groupes de stimuli. Nous nous attendons donc à observer la constitution de 4 groupes de produits au maximum. Cependant, certains sujets peuvent décider de considérer des stimuli « moyens », situés par conséquent au centre de la nappe, constituant ainsi un cinquième groupe. » Nous avons donc fixé le nombre des étapes à 4 pour permettre d'obtenir une partition en 5 groupes de stimuli.

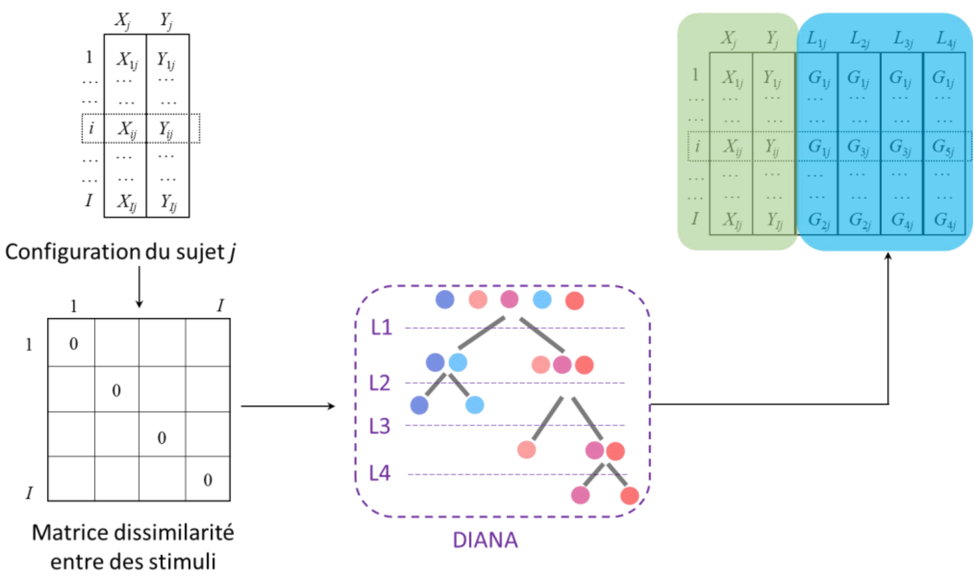


Figure D-4 Données simulées sous contrainte comportementale en effectuant l'algorithme divisif (DIANA) sur des données de Napping®

Les résultats fournis dans le Chapitre 2 permettent d'identifier quatre types de comportement du sujet. Ces résultats sont directement liés au modèle comportemental utilisé et à l'algorithme divisif. Ils ne permettent pas de mettre en évidence des processus cognitifs plus complexes où le sujet change d'avis au cours de la tâche par exemple. Néanmoins ces simulations nous ont permis de réfléchir à notre solution de recueil et d'analyse, en attendant d'obtenir des données expérimentales.

Quatrième point – De l'apport des données de digit-tracking dans la compréhension du Napping®

Les données utilisées sont issues d'une expérimentation, où on demande à des sujets de positionner des vidéos qui correspondent à des publicités de parfums pour hommes. Pour illustrer la représentation du processus cognitif au cours du temps, nous avons considéré 4 sujets différents. Rappelons que les représentations de ces 4 sujets sont interprétées conjointement avec la représentation des stimuli où : la première dimension oppose les publicités associées aux mots «nature, sauvage» d'une part, «jeune, festif» d'autre part ; la deuxième dimension oppose les publicités associées au concept «séduction», aux autres publicités. D'après la représentation des sujets, on peut conclure que : le sujet 42 a séparé les stimuli en fonction de la Dim 1, les sujets 32 et 25 ont séparé les stimuli en fonction de la Dim 2, et le sujet 8 n'a identifié aucune des deux premières dimensions. Figure 3.5 représente le processus cognitif au cours du temps des 4 sujets. Alors que les sujets 32 et 25 semblent proches du point de vue de leur nappe, ils se révèlent très différent du point de vue de leur processus cognitif : la dimension sensorielle perçue par le sujet 32 est restée la même tout au long de l'expérience, alors que pour le sujet 25, elle a évolué.

L'intégration du comportement du sujet dans l'analyse des données de Napping® a été motivée par l'étude de la stabilité des dimensions fournies par un sujet. Plus détaillé, les données digit-tracking apporte dans la compréhension du Napping® selon 3 angles. Un premier associé à une meilleure compréhension de l'espace des stimuli, un second à une meilleure compréhension des sujets, enfin un troisième associé à une meilleure compréhension de la tâche de Napping®.

Doit-on considérer que la deuxième dimension qui oppose les vidéos associées à la séduction aux autres est plus stable pour le sujet 32 qu'elle ne l'est pour le sujet 25 ? Deux points de vue s'affrontent :

- Le premier consiste à dire que la dimension 2 est extrêmement stable pour le sujet 25 parce qu'il a justement positionné les vidéos selon cette dimension seulement alors qu'il avait perçu la première dimension.
- Le second consiste à mettre au même niveau de stabilité la dimension 2 pour les sujets 32 et 25, mais en renforçant la stabilité de la dimension 1 au niveau du panel, cette dernière ayant été perçue par le sujet 25.

C'est ce dernier point de vue que nous adopterons. Ainsi, nous relativiserons la stabilité d'une dimension au niveau du panel en fonction du nombre de fois qu'elle aura été perçue au cours de la tâche de Napping®. Une dimension sera d'autant plus stable qu'elle aura été perçue un grand nombre de fois (cf. Figure D-5).

Dans notre donnée (des vidéos), quatre autres sujets ont le même processus cognitif que le sujet 25. Il s'agit des sujets 10, 37, 61, et 70 (cf. Figure D-6). Le fait qu'ils aient perçu la première dimension va renforcer la stabilité de cette dernière dans notre interprétation.

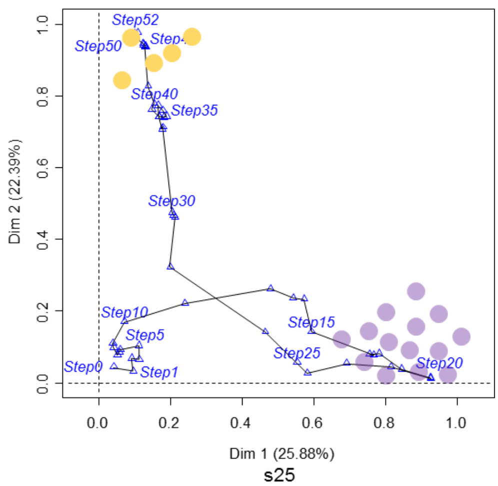


Figure D-5 Illustration de la stabilité d'une dimension au niveau du panel. Une dimension sera d'autant plus stable qu'elle aura été perçue un grand nombre de fois

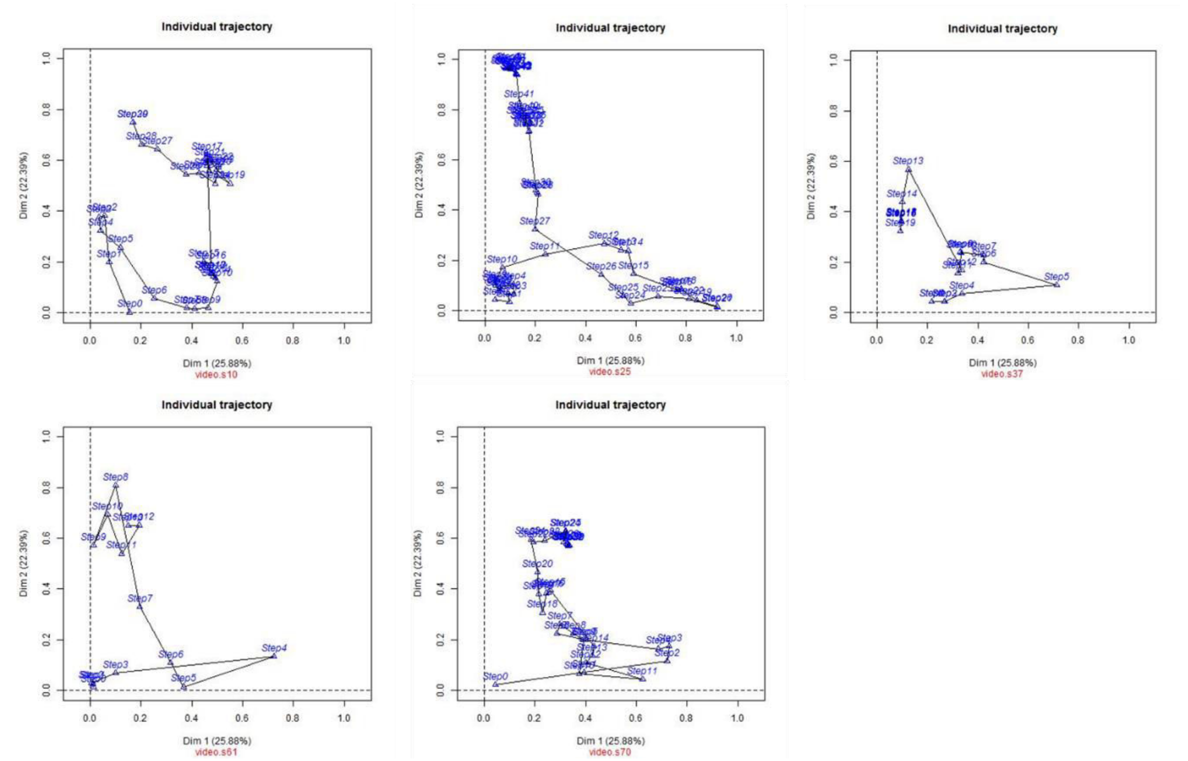


Figure D-6 Quatre autres sujets ont le même processus cognitif que le sujet 25

Au-delà de l'étude de la stabilité des dimensions qui nous permet de mieux interpréter l'espace des stimuli, on peut analyser les données de digit-tracking pour mieux comprendre le processus cognitif des sujets. Ici également, deux points de vue s'offrent à nous :

- Un premier qui vise à construire une typologie des sujets en fonction de leur seul comportement lors de la tâche de Napping®. Pour cette typologie, deux sujets seront d'autant plus proches qu'ils auront des processus cognitifs proches indépendamment de l'ordre des dimensions. Ainsi nous aurons par exemple, des sujets qui ne changent pas d'avis dans un même groupe, dans un autre groupe au contraire, des sujets qui changent d'avis.
- Le second point de vue, prend en compte à la fois le comportement et les dimensions obtenues par le Napping®. Parmi les sujets qui ne changent pas d'avis, nous distinguerons ceux qui ont perçu la première dimension de ceux qui ont perçu la seconde dimension.

Enfin, on pourra également utiliser les données de digit-tracking pour mieux comprendre la tâche de Napping® en tant que telle. Les stimuli lors de la tâche de Napping® sont présentés au sujet dans leur ensemble, suivant un ordre aléatoire. L'ordre de présentation y-a-t-il un impact sur la configuration finale ? Il est difficile d'apporter une réponse tranchée à cette question. Si on regarde les sujets semblables au sujet 25, il semblerait qu'ils aient abouti aux mêmes conclusions, suivant le même processus cognitif à partir d'un ordre de présentation différent. Cet élément de réponse semble aller en faveur d'un ordre de présentation aléatoire.

Cinquième point – Que peut-on attendre des données de digit-tracking ?

Qu'avons-nous fait pendant trois ans ? Nous avons réfléchi au Napping®. Nous avons essayé de mieux comprendre ce recueil de données sensorielles proposé voilà déjà 10 ans par Jérôme Pagès.

Cette compréhension s'est faite à travers différents angles d'attaque, différentes questions. Notamment : Quels sont les objectifs du Napping® ? Quelle est la nature des données recueillies par le Napping® ? C'est à partir de cette compréhension que nous avons identifié une question centrale : Quelles données supplémentaires faut-il recueillir pour mieux comprendre les données de Napping® ?

Après avoir souligné l'intérêt de recueillir des données comportementales pour une meilleure compréhension des données de Napping® nous avons créé Holos, l'environnement de recueil de données appelées Digit-Tracking. Nous avons développé une solution d'analyse que nous avons intégré dans le package SensoMineR. Néanmoins des questions restent en suspens.

La première concerne la façon dont nous avons représenté graphiquement les processus cognitifs associés au sujet. Cette représentation utilise les données de tracking comme une information supplémentaire. Elle nous permet de mieux appréhender la représentation consensuelle calculée sur les images finales fournies par les sujets. Sommes-nous allés suffisamment loin dans l'intégration des données de tracking ? Autrement dit faut-il leur accorder le statut de variable active ? Des essais ont déjà été entrepris ils se révèlent encourageant. Cette prise en compte de l'ensemble des processus cognitifs en tant qu'éléments actifs permettrait de relativiser directement l'importance d'une dimension par rapport à une autre dès lors qu'elles ont été perçues au cours du temps. Cette prise en compte du processus cognitif en tant qu'élément actif de l'analyse correspondrait d'ailleurs finalement à ce qui est déjà fait dans le traitement des données de Napping® catégorisé et dans celui des données de tri hiérarchique. L'adaptation aux données de tracking bien plus complexes semble faisable en théorie.

La seconde question concerne l'étude du processus cognitif associé aux sujets. Quelle distance utiliser pour établir une typologie sur les sujets ? Cette question doit être raisonnée rapport aux applications qui reste encore à identifier clairement. La réponse à cette question semble être multidisciplinaire. Elle relève à la fois de la psychologie, du sensoriel, et de la statistique. Dans quelle mesure peut-on exploiter des sujets classés selon leur processus cognitif ? Dans quelle mesure peut-on réaliser une classification de sujet en fonction de leur processus cognitif ?

(This page is left blank)

Conclusion

Ce manuscrit représente une contribution à une meilleure compréhension du Napping® du point de vue du recueil de données d'une part, de l'analyse de données d'autre part. L'apport principal de ce travail de recherche réside dans l'intégration d'une information supplémentaire inhérente au Napping®, à savoir le processus cognitif du sujet.

L'introduction de ce manuscrit retrace très brièvement l'histoire du Napping®, de sa création à ses applications les plus récentes. Comme nous l'avons rappelé, le Napping® s'est développé au cours de la dernière décennie, à une période où l'intégration des consommateurs dans le processus de développement produit est devenue une pratique de plus en plus répandue. C'est précisément parce que la tâche de Napping® pouvait être réalisée par des consommateurs, et parce qu'elle fournissait des résultats en apparence comparables à ceux d'un profil sensoriel quantitatif souvent plus longs et plus couteux à obtenir, que le Napping® a connu un engouement certain, auprès des industriels notamment.

Le chapitre 1 de ce manuscrit constitue la base de notre contribution à une meilleure compréhension du Napping® : nous y expliquons la nature intrinsèque des données recueillies lors de la tâche de Napping®. En particulier, nous insistons sur deux points.

Premièrement, les données de Napping® sont de nature holistique, et non analytique. Par conséquent, il ne nous semble pas raisonnable d'évaluer les "performances" du Napping® en se limitant à la seule comparaison des résultats que cette méthode fournit à ceux d'une méthode analytique. L'intérêt principal du Napping® est que cette méthode permet l'émergence des principaux axes de variabilité d'un espace produit perçus par un ensemble de sujets, souvent naïfs.

Deuxièmement, la prise en compte du comportement du sujet dans la façon dont les données sont intégrées puis analysées est cruciale. Cette prise en compte nous a conduits au Napping catégorisé, une évolution naturelle du Napping®. C'est en observant les sujets lors de la phase de verbalisation que l'on s'est aperçu qu'ils verbalisaient au niveau des groupes de stimuli et non pas au niveau de chaque stimulus. C'est ainsi qu'une variable qualitative de groupe, correspondant aux données de verbalisation a été rajoutée aux coordonnées des stimuli recueillies originellement. Cette variable additionnelle est indispensable à la compréhension de la représentation des stimuli obtenue sur la base des coordonnées fournies par les sujets.

Réfléchir à la nature des données de Napping® et la comprendre, nous a permis de prendre la pleine mesure de l'importance du sujet dans le recueil des données. La compréhension d'un ensemble de stimuli ne passe plus simplement par la compréhension des seuls stimuli, elle doit pleinement intégrer le rôle majeur que prend le sujet au cœur de l'évaluation sensorielle : contrairement au profil sensoriel quantitatif classique de type QDA®, où le sujet (le plus souvent, dénommé juge) est entrainé à évaluer à travers une grille de lecture fixée a priori, il est expressément demandé au sujet lors d'une tâche de Napping® d'évaluer selon sa propre grille de lecture. On comprend pourquoi l'utilisation des seules coordonnées des stimuli ne peut suffire à une compréhension fine des résultats, et qu'il faut au-delà des distances inter-stimuli amener de l'information sur le sujet.

Le chapitre 2 nous a confortés dans notre direction de recherche, à savoir la nécessité de mesurer et d'intégrer le comportement des sujets afin de mieux comprendre les résultats fournis par le Napping®. C'est dans ce chapitre qu'est véritablement apparue la notion de processus cognitif d'un sujet lors de la tâche de Napping® : à défaut de le mesurer, nous avons proposé une façon de le modéliser. Ce modèle a été construit sur la base d'une hypothèse forte, mais réaliste : on prête aux sujets une stratégie dite descendante (top-down approach) lorsqu'ils font face à un grand ensemble de stimuli (ce comportement a été majoritairement observé en pratique). À partir des distances inter-stimuli fournies par chaque sujet et de l'algorithme de classification divisif DIANA, nous avons pu obtenir des représentations graphiques des processus cognitifs des sujets. L'arbre de classification descendant hiérarchique fourni par DIANA pour chaque sujet est recodé sous forme de variables qualitatives. Cette dernière information est ensuite utilisée comme une information supplémentaire dans l'AFM du tableau des coordonnées des stimuli fournies par l'ensemble des sujets.

Ces représentations graphiques nous ont permis de visualiser la multiplicité des processus cognitifs modélisés. C'est cette diversité des processus qui nous a amené naturellement à la question de leur classification, question à laquelle nous n'avons apporté qu'une réponse partielle mais pragmatique. Ces représentations nous ont surtout permis de réaliser qu'il était indispensable de mesurer l'ensemble des processus cognitifs pour les comprendre, et que leur apparente multiplicité était en fait sous-estimée. Une même configuration finale pouvant être obtenue en passant par différentes étapes, il semble illusoire de comprendre la notion de processus sur la base de sa réalisation et d'une hypothèse comportementale. L'interaction entre les stimuli et le sujet, et par conséquent les multiples clés d'entrée pour aborder un ensemble de stimuli, la possibilité d'en changer au cours du temps, se révèlent être des éléments indispensables à la compréhension du processus cognitif d'un sujet.

C'est finalement dans le chapitre 3 que nous proposons le digit-tracking, une nouvelle méthode de recueil de données, ou plutôt une façon de recueillir des données complémentaires à celles originellement recueillies lors d'une tâche de Napping®. Dans ce même chapitre, nous proposons également une façon d'analyser ces données à travers des représentations graphiques originales que nous interprétons. De la visualisation des processus cognitifs, il est possible de tirer des enseignements sur la représentation de l'espace des stimuli obtenue à partir de l'ensemble des sujets. La compréhension de l'évolution des dimensions sensorielles individuelles, nous permet entre autres d'appréhender la stabilité des dimensions sensorielles perçues par l'ensemble des sujets, et de mieux évaluer l'importance relative d'une dimension sensorielle par rapport à une autre. La méthode du digit-tracking que nous proposons repose sur une nouvelle technologie basée sur l'utilisation d'une tablette numérique tactile couplée à un nouvel environnement de travail que nous avons nommé Holos.

Le chapitre 4 constitue la partie émergée de nos travaux de recherche. Nous y présentons l'environnement Holos, ainsi que les fonctions qui permettent le traitement des données recueillies par digit-tracking, que l'on peut finalement qualifier de "données de Napping®

temporelles". Au-delà de sa fonction première qui est de pouvoir recueillir des données pour tous types de stimuli de façon pratique (recueil en tant que tel et stockage des données), Holos a été pensé comme un environnement collaboratif permettant à des chercheurs de partager leurs ressources, en partie ou en totalité.

Ce travail de recherche a été l'occasion pour nous de réfléchir au Napping®, une méthode proposée il y a 10 ans de cela par Jérôme Pagès. L'enseignement principal que nous avons tiré de ces réflexions est qu'il faut comprendre en profondeur la nature d'une mesure pour l'exploiter pleinement. La mesure, qui est par définition l'action d'évaluer une grandeur, doit être, dans le cas du Napping®, étudiée du début à la fin de la tâche, et l'on ne peut se contenter de la seule configuration finale individuelle pour comprendre la configuration finale basée sur l'ensemble des sujets. Des données de Napping® originelles nous sommes finalement parvenus aux données de Napping® temporelles : le Napping® est non seulement une approche holistique mais également et peut être surtout une approche dynamique.

Cette nouvelle dimension temporelle ouvre de nouvelles voies de développement et de recherche, puisqu'elle nous amène à intégrer un nouveau concept, celui du processus cognitif.

En ce qui concerne le développement, proposer un nouveau point de vue sur une tâche maintenant couramment utilisée, implique de parfaire l'outil technologique qui permet le recueil des données associées à ce nouveau point de vue. Afin que l'outil développé soit effectivement utilisé, nous devrons veiller à ce que les utilisateurs puissent travailler en toute sécurité. Ce problème est maintenant assez classique dans le développement d'applications dites *cloud application*s, qui permettent le stockage des données sur serveur.

En ce qui concerne la recherche, d'un point de vue statistique se pose la question de la classification des processus cognitifs. Il s'agit d'un problème de classification de trajectoires multidimensionnelles qu'il faut adapter au cas particulier des données de Napping®. En pratique, nous nous poserons également le problème de la mise en relation des résultats issus de cette classification avec des données extérieures portant sur le sujet. Ce problème en apparence classique, reste ouvert d'un point de vue psychologie cognitive, puisqu'il s'agit de trouver la mesure qui permettrait de comprendre la perception dynamique qu'un sujet peut avoir d'un ensemble de stimuli.

(This page is left blank)

References

Ares, G., de Saldamando, L., Vidal, L., Antúnez, L., Giménez, A., & Varela, P. (2013). Polarized Projective Mapping: Comparison with Polarized Sensory Positioning approaches. *Food Quality and Preference*, *28*(2), 510–518. doi:10.1016/j.foodqual.2013.01.003

Bak, A., & Bartlomowicz, T. (2012). Conjoint analysis method and its implementation in conjoint R package. Retrieved from http://cran.r-project.org/web/packages/conjoint/index.html

Barcenas, P., Pérez Elortondo, F. J., & Albisu, M. (2004). Projective mapping in sensory analysis of ewes milk cheeses: A study on consumers and trained panel performance. *Food Research International*, *37*(7), 723–729. doi:10.1016/j.foodres.2004.02.015

Barrett, L. F. (1998). Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus. *Cognition & Emotion*, *12*(4), 579–599. doi:10.1080/026999398379574

Bécue-Bertaut, M., & Lê, S. (2011). Analysis Of Multilingual Labeled Sorting Tasks: Application To A Cross-Cultural Study In Wine Industry. *Journal of Sensory Studies*, *26*(5), 299–310. doi:10.1111/j.1745-459X.2011.00345.x

Berland, A., Gaillard, P., Guidetti, M., & Barone, P. (2015). Perception of Everyday Sounds: A Developmental Study of a Free Sorting Task. *PloS One*. doi:10.1371/journal.pone.0115557

Cadena, R. S., Caimi, D., Jaunarena, I., Lorenzo, I., Vidal, L., Ares, G., … Giménez, A. (2014). Comparison of rapid sensory characterization methodologies for the development of functional yogurts. *Food Research International*, *64*, 446–455. doi:10.1016/j.foodres.2014.07.027

Cadoret, M. (2010). *Analyse factorielle de données de catégorisation application aux données sensorielles*. Agrocampus.

Cadoret, M., Lê, S., & Pagès, J. (2009). A factorial approach for sorting task data (FAST). *Food Quality and Preference*, *20*(6), 410–417. doi:10.1016/j.foodqual.2009.02.010

Cadoret, M., Lê, S., & Pagès, J. (2011). Statistical analysis of hierarchical sorting data. *Journal of Sensory Studies*, *26*(2), 96–105. doi:10.1111/j.1745-459X.2010.00326.x

Carroll, J. D. (1968). Generalization of canonical correlation analysis to three or more sets of variables. *Proceedings of the 76th Annual Convention of the American Psychological Association*, *3*, 227–228.

Chollet, S., Lelièvre, M., Abdi, H., & Valentin, D. (2011). Sort and beer: everything you wanted to know about the sorting task but did not dare to ask. *Food Quality and Preference*, *22*(6), 507–520. doi:10.1016/j.foodqual.2011.02.004

Chrea, C., Grandjean, D., Delplanque, S., Cayeux, I., Le Calvé, B., Aymard, L., … Scherer, K. R. (2009). Mapping the semantic space for the subjective experience of emotional responses to odors. *Chemical Senses*, *34*(1), 49–62. doi:10.1093/chemse/bjn052

Costa, A. I. A., & Jongen, W. M. F. (2006). New insights into consumer-led food product development. *Trends in Food Science & Technology*, *17*(8), 457–465. doi:10.1016/j.tifs.2006.02.003

Courcoux, P., Qannari, E. M., & Faye, P. (2015). Free sorting as a sensory profiling technique for product development. In J. Delarue, B. Lawlor, & M. Rogeaux (Eds.), *Woodhead Publishing Series in Food Science, Technology and Nutrition* (pp. 153–185). United Kingdom: Woodhead

Publishing. doi:http://dx.doi.org/10.1533/9781782422587.1.3

Cruz, a. G., Cadena, R. S., Castro, W. F., Esmerino, E. a., Rodrigues, J. B., Gaze, L., … Bolini, H. M. a. (2013). Consumer perception of probiotic yogurt: Performance of check all that apply (CATA), projective mapping, sorting and intensity scale. *Food Research International*, *54*(1), 601–610. doi:10.1016/j.foodres.2013.07.056

Dadzie, B. K., & Orchard, J. E. (1997). *Routine post-harvest screening of banana/plantain hybrids: criteria and methods* (INIBAP Tec.). CIRPAC.

Darwin, C. (1872). *The expression of the emotions in man and animals*. (F. Darwin, Ed.) (3rd ed.). Oxford University Press. doi:10.1097/00000441-195610000-00024

Dehlholm, C. (2014). Projective mapping and Napping. In P. Valera & G. Ares (Eds.), *Novel techniques in sensory characterization and consumer profiling* (pp. 229–253). CRC Press.

Dehlholm, C., Brockhoff, P. B., Meinert, L., Aaslyng, M. D., & Bredie, W. L. P. (2012). Rapid descriptive sensory methods – Comparison of Free Multiple Sorting, Partial Napping, Napping, Flash Profiling and conventional profiling. *Food Quality and Preference*, *26*(2), 267–277. doi:10.1016/j.foodqual.2012.02.012

Delarue, J. (2015). The use of rapid sensory methods in R&D and research: an introduction. In J. Delarue, B. Lawlor, & M. Rogeaux (Eds.), *Rapid sensory profiling techniques and related methods* (pp. 3–25). United Kingdom: Woodhead Publishing. doi:http://dx.doi.org/10.1533/9781782422587.1.3

Donoghue, S. (2000). Projective techniques in consumer research. *Journal of Family Ecology and Consumer Sciences*, *28*, 47–53.

Ekman, P., & Cordaro, D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review*, *3*(4), 364–370. doi:10.1177/1754073911410740

Escofier, B., & Pagès, J. (1994). Multiple factor analysis (AFMULT package). *Computational Statistics & Data Analysis*, *18*(1), 121–140. doi:10.1016/0167-9473(94)90135-X

Escofier, B., & Pagès, J. (2008). *Analyses factorielles simples et multiples: Objectifs, méthodes et interprétation* (Fourth.). Paris: Dunod.

Faye, P. (2012). *Étude des différences de perception entre sujets - approche méthodologique et application à une épreuve de catégorisation*. Oniris.

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, *18*(12), 1050–7. doi:10.1111/j.1467-9280.2007.02024.x

Gaillard, P. (2009). Laissez-nous trier ! TCL-LabX et les tâches de catégorisation libre de sons. In D. Dubois (Ed.), *Le Sentir et le Dire : Concepts et méthodes en psychologie et linguistique cognitives* (pp. 189–210). Paris: L'harmattan.

Giacalone, D., Bredie, W. L. P., & Frøst, M. B. (2013). "All-In-One Test" (AI1): A rapid and easily applicable approach to consumer product testing. *Food Quality and Preference*, *27*(2), 108–119. doi:10.1016/j.foodqual.2012.09.011

Giacalone, D., Ribeiro, L. M., & Frøst, M. B. (2013). Consumer-Based Product Profiling: Application of Partial Napping® for Sensory Characterization of Specialty Beers by Novices and Experts. *Journal of Food Products Marketing*, *19*(3), 201–218. doi:10.1080/10454446.2013.797946

Grunert, K. G. (2002). Current issues in the understanding of consumer food choice. *Trends in Food Science & Technology*, *13*(8), 275–285. doi:10.1016/S0924-2244(02)00137-1

Hopfer, H., & Heymann, H. (2013). A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. *Food Quality and Preference*, *28*(1), 164–181. doi:10.1016/j.foodqual.2012.08.017

Husson, F., Josse, J., Lê, S., & Mazet, J. (2014). FactoMineR: multivariate exploratory data analysis and data mining with R. Retrieved from http://cran.r-project.org/package=FactoMineR

Husson, F., Le Dien, S., & Pagès, J. (2001). Which value can be granted to sensory profiles given by consumers? Methodology and results. *Food Quality and Preference*, *12*(5-7), 291–296. doi:10.1016/S0950-3293(01)00014-3

Husson, F., Lê, S., & Pagès, J. (2011). *Exploratory multivariate analysis by example using R*. London: Chapman & Hall/CRC.

Izard, C. (2011). Forms and functions of emotions: Matters of emotion–cognition interactions. *Emotion Review*, *3*(4), 371–378. doi:10.1177/1754073911410737

Jourdan, S., Morvant, S., Lê, S., & Pagès, J. (2014). Liking sensory space and marketing space: example of ten luxury men perfumes. *Poster Presented at Sensometrics*.

Kennedy, J. (2010). Evaluation of replicated projective mapping of granola bars. *Journal of Sensory Studies*, *25*(5), 672–684. doi:10.1111/j.1745-459X.2010.00302.x

Kennedy, J., & Heymann, H. (2009). Projective Mapping and Descriptive Analysis of Milk and Dark Chocolates. *Journal of Sensory Studies*, *24*(2), 220–233. doi:10.1111/j.1745-459X.2008.00204.x

Kim, Y. K., Jombart, L., Valentin, D., & Kim, K. O. (2013). A cross-cultural study using Napping®: Do Korean and French consumers perceive various green tea products differently? *Food Research International*, *53*(1), 534–542. doi:10.1016/j.foodres.2013.05.015

King, M. C., Cliff, M. a., King, M. C., Hall, J. W., Cliff, M. a, & Hall, J. W. (1998). Comparison of projective mapping and sorting data collection and multivariate methodologies for identification of similarity-of-use of snack bars. *Journal of Sensory Studies*, *13*(3), 347–358. doi:10.1111/j.1745-459X.1998.tb00094.x

Laros, F. J. M., & Steenkamp, J.-B. E. M. (2005). Emotions in consumer behavior: a hierarchical approach. *Journal of Business Research*, *58*(10), 1437–1445. doi:10.1016/j.jbusres.2003.09.013

Larsen, R. J., & Diener, E. (1992). Promises and Problems With the Circurnplex Model of Emotion. *Review of Personality and Social Psychology*, *13*, 25–57.

Le Dien, S., & Pagès, J. (2003). Hierarchical Multiple Factor Analysis: application to the comparison of sensory profiles. *Food Quality and Preference*, *14*(5-6), 397–403. doi:10.1016/S0950-3293(03)00027-2

Lê, S., & Husson, F. (2008). Sensominer: A package for sensory data analysis. *Journal of Sensory Studies*, *23*(1), 14–25. doi:10.1111/j.1745-459X.2007.00137.x

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 1–18.

Lê, S., Lê, T. M., & Cadoret, M. (2015). Napping and sorted Napping as a sensory profiling technique. In J. Delarue, B. Lawlor, & M. Rogeaux (Eds.), *Rapid sensory profiling techniques and related methods* (pp. 197–213). United Kingdom: Woodhead Publishing. doi:http://dx.doi.org/10.1533/9781782422587.2.197

Lê, T. M., & Lê, S. (2014). Holos: a collaborative environment for holistic approaches. *Poster Presented at Sensometrics*. Chicago, IL, USA.

Lebart, L., Morineau, A., & Warwick, K. M. (1984). *Multivariate descriptive statistical analysis: correspondence analysis and related techniques for large matrices*. New York, NY: Wiley.

Levenson, R. (2011). Basic emotion questions. *Emotion Review*, *3*(4), 379–386. doi:10.1177/1754073911410743

Lorenzen, T., & Anderson, V. (1993). *Design of experiments: a no-name approach*. New York, NY: CRC Press.

Louw, L., Malherbe, S., Naes, T., Lambrechts, M., van Rensburg, P., & Nieuwoudt, H. (2013). Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products. *Food Quality and Preference*, *30*(2), 192–201. doi:10.1016/j.foodqual.2013.05.016

MacFie, H. (Ed.). (2007). *Consumer-led food product development*. Cambridge: Woodhead Publishing.

MacNaughton-Smith, P., Williams, W. T., Dale, M. B., & Mockett, L. G. (1964). Dissimilarity Analysis: a new Technique of Hierarchical Sub-division. *Nature*, *202*(4936), 1034–1035. doi:10.1038/2021034a0

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2014). Cluster: cluster analysis basics and extensions. Retrieved from http://cran.r-project.org/web/packages/cluster/index.html

Moskowitz, H., & Krieger, B. (1993). What sensory characteristics drive product quality? An assessment of individual differences. *Journal of Sensory Studies*, *8*(4), 271–282. doi:10.1111/j.1745-459X.1993.tb00219.x

Moskowitz, H. R. (1995). Food quality: Conceptual and sensory aspects. *Food Quality and Preference*, *6*(3), 157–162. doi:10.1016/0950-3293(94)00030-Y

Munoz, A. M., & Civille, G. V. (1992). The spectrum descriptive analysis method. In R. Hootman (Ed.), *Manual on Descriptive Analysis Testing for Sensory Evaluation* (pp. 22–34). 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959: ASTM International. doi:10.1520/MNL13-EB

Nestrud, M. a., & Lawless, H. T. (2008). Perceptual mapping of citrus juices using projective mapping and profiling data from culinary professionals and consumers. *Food Quality and Preference*, *19*, 431–438. doi:10.1016/j.foodqual.2008.01.001

Nestrud, M. A., & Lawless, H. T. (2010). Perceptual mapping of apples and cheeses using projective mapping and sorting. *Journal of Sensory Studies*, *25*(3), 390–405. doi:10.1111/j.1745-459X.2009.00266.x

Nestrud, M. A., & Lawless, H. T. (2011). Recovery of subsampled dimensions and configurations derived from napping data by MFA and MDS. *Attention, Perception & Psychophysics*, *73*(4), 1266–1278. doi:10.3758/s13414-011-0091-0

Niedenthal, P., & Halberstadt, J. (2003). Top-down influences in social perception. *European Review of Social Psychology*, *14*, 49–76. doi:10.1080/10463280340000018

Pagès, J. (2004). Multiple factor analysis: Main features and application to sensory data. *Revista Colombiana de Estadística*, *27*(1), 1–26.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire Valley. *Food Quality and Preference*, *16*(7), 642–649. doi:10.1016/j.foodqual.2005.01.006

Pagès, J. (2013). *Analyse factorielle multiple avec R* (1st ed., Vol. 4). EDP Sciences.

Pagès, J. (2015). *Multiple factor analysis by example using R*. Chapman & Hall/CRC.

Pagès, J., Cadoret, M., & Lê, S. (2010). The sorted napping: a new holistic approach in sensory evaluation. *Journal of Sensory Studies*, *25*(5), 637–658. doi:10.1111/j.1745-459X.2010.00292.x

Pagliuca, M., & Scarpato, D. (2011). Food quality, consumer perception and preferences: an analysis on olive oil. *Electronic Journal of Applied Statistical …*, *4*(2), 215–226. doi:10.1285/i20705948v4n2p215

Panksepp, J., & Watt, D. (2011). What is Basic about Basic Emotions? Lasting Lessons from Affective Neuroscience. *Emotion Review*, *3*(4), 387–396. doi:10.1177/1754073911410741

Perrin, L., & Pagès, J. (2009). Construction of a product space from the ultra-flash profiling method: Application to 10 red wines from the loire valley. *Journal of Sensory Studies*, *24*(3), 372–395. doi:10.1111/j.1745-459X.2009.00216.x

Perrin, L., Symoneaux, R., Maître, I., Asselin, C., Jourjon, F., & Pagès, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: Case of ten wines from Loire valley. *Food Quality and Preference*, *19*(1), 1–11. doi:10.1016/j.foodqual.2007.06.005

Pineau, N., & Schlich, P. (2015). Temporal dominance of sensations (TDS) as a sensory profiling technique. In J. Delarue, B. Lawlor, & M. Rogeaux (Eds.), *Rapid sensory profiling techniques and related methods* (pp. 269–306). United Kingdom: Woodhead Publishing.

R Core Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.r-project.org/

Rao, V. R. (2014). *Applied Conjoint Analysis*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-87753-0

Ratings, S. (2012). 2 . 3 Partial Order Scoring (POSCOR), (d), 10–11.

Reinbach, H. C., Giacalone, D., Ribeiro, L. M., Bredie, W. L. P., & Frøst, M. B. (2014). Comparison of three sensory profiling methods based on consumer perception: CATA, CATA with intensity and Napping®. *Food Quality and Preference*, *32*, 160–166. doi:10.1016/j.foodqual.2013.02.004

Risvik, E., McEwan, J. A., Colwill, J. S., Rogersa, R., & Lyonb, D. H. (1994). Projective mapping: a tool for sensory analysis and consumer research. *Food Quality and Preference*, *5*, 263–269.

Risvik, E., McEwan, J. a., & Rødbotten, M. (1997). Evaluation of sensory profiling and projective mapping data. *Food Quality and Preference*, *8*(1), 63–71. doi:10.1016/S0950-3293(96)00016-X

Robitza, W., Pitrey, Y., & Hlavacs, H. (2013). The NappingPlayer — Projective Mapping Experiments on Android Tablets. In *PQS - Fourth International Workshop on Perceptual Quality of Systems* (pp. 1–6).

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Santos, B. a., Pollonio, M. a R., Cruz, a. G., Messias, V. C., Monteiro, R. a., Oliveira, T. L. C., … Bolini, H. M. a. (2013). Ultra-flash profile and projective mapping for describing sensory attributes of prebiotic mortadellas. *Food Research International*, *54*(2), 1705–1711. doi:10.1016/j.foodres.2013.09.022

Santosa, M., Abdi, H., & Guinard, J. X. (2010). A modified sorting task to investigate consumer perceptions of extra virgin olive oils. *Food Quality and Preference*, *21*(7), 881–892. doi:10.1016/j.foodqual.2010.05.011

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: further exploration of a prototype approach. *Journal of Personality and Social Psychology*, *52*(6), 1061–1086.

Steinman, R. B. (2009). Projective techniques in consumer research, *5*(5), 37–45.

Stone, H., Sidel, J., Oliver, S., Woosley, A., & Singleton, R. C. (1974). Sensory evaluation by quatitative descriptive analysis. *Food Technology*, *28*, 24–34.

Storm, C., & Storm, T. (1987). A taxonomic study of the vocabulary of emotions. *Journal of Personality and Social Psychology*, *53*(4), 805–816.

Thomson, D. M. H., & Crocker, C. (2013). A data-driven classification of feelings. *Food Quality and Preference*, *27*(2), 137–152. doi:10.1016/j.foodqual.2012.09.002

Torri, L., Dinnella, C., Recchia, A., Naes, T., Tuorila, H., & Monteleone, E. (2013). Projective Mapping for interpreting wine aroma differences as perceived by naïve and experienced assessors. *Food Quality and Preference*, *29*(1), 6–15. doi:10.1016/j.foodqual.2013.01.006

Tracy, J. L., & Randles, D. (2011). Four Models of Basic Emotions: A Review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review*, *3*(4), 397–405. doi:10.1177/1754073911410747

Tsuji, M., Shimokawa, T., & Okada, A. (2010). Three-way scaling and clustering approach to musical structural analysis. In H. Locarek-Junge & C. Weihs (Eds.), *Classification as a tool for research* (pp. 767–774). Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-10745-0_85

Valentin, D., Chollet, S., Lelièvre, M., & Abdi, H. (2012a). Quick and dirty but still pretty good: A review of new descriptive methods in food science. *International Journal of Food Science and Technology*, *47*, 1563–1578. doi:10.1111/j.1365-2621.2012.03022.x

Valentin, D., Chollet, S., Lelièvre, M., & Abdi, H. (2012b). Quick and dirty but still pretty good: A review of new descriptive methods in food science. *International Journal of Food Science and Technology*, *47*(8), 1563–1578. doi:10.1111/j.1365-2621.2012.03022.x

Veinand, B., Godefroy, C., Adam, C., & Delarue, J. (2011). Highlight of important product characteristics for consumers. Comparison of three sensory descriptive methods performed by consumers. *Food Quality and Preference*, *22*(5), 474–485. doi:10.1016/j.foodqual.2011.02.011

Vidal, L., Cadena, R. S., Antúnez, L., Giménez, A., Varela, P., & Ares, G. (2014). Stability of sample configurations from projective mapping: How many consumers are necessary? *Food Quality and Preference*, *34*, 79–87. doi:10.1016/j.foodqual.2013.12.006

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*, 219–235. doi:10.1037/0033-2909.98.2.219

Whitley, L. (2010). Consumer-led food product development. *International Journal of Dairy Technology*, *63*(2), 301–302. doi:10.1111/j.1471-0307.2009.00558.x

Wickham, H. (2009). *ggplot2*. *Wiley Interdisciplinary Reviews: Computational …*. New York, NY: Springer New York. doi:10.1007/978-0-387-98141-3

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*(5), 558–568. doi:10.1002/jsfa.2740350513

PhD project activities

Accepted publications

Lê, M.T., Brard, M., & Lê, S. (2016, accepted). Holos: A collaborative environment for similarity-based holistic approaches, Behaviour Research Methods.

Lê, S., Lê, M.T., & Cadoret, M. (2015). Napping® and sorted napping as a sensory profiling technique. In J. Delarue, B. Lawlor, & M. Rogeaux (Eds.), Rapid sensory profiling techniques and related methods (pp. 197–213). United Kingdom: Woodhead Publishing. doi: 10.1533/9781782422587.2.197

Lê, M.T., Husson, F., & Lê, S. (2015). Digit-tracking: interpreting the evolution over time of sensory dimensions of an individual product space issued from Napping® and sorted Napping, Food Quality and Preference. doi: 10.1016/j.foodqual.2015.07.002

Lê, M.T., Lê, S., & Nguyen, H.D. (2014). Assessing consumer-perceived food quality using conjoint analysis. Journal of Science development, Journal of Science and Technology Development, 17-27, Vol.17, ISSN: 1859-0128.

Submitted publication

Lê, M.T., & Lê, S. (2014). A Napping® based methodology to quickly obtain a model of emotions, Behaviour Research Methods.

Software

Lê, M.T., & Lê, S. (2014). Holos: a collaborative environment for collecting data from holistic methods, http://napping.agrocampus-ouest.fr/

Lê, M.T., & Lê, S. (2015). The R functions (integrated in SensoMineR) for analysing temporal data issued from Napping® method.

Oral presentations at international conferences

Lê, M.T., & Lê, S. (2014). Holos: an environment for studying the evolution of sensory dimensions of individual product space based on Holistic methods. EuroSense2014, Copenhagen, Denmark.

Lê, M.T., & Lê, S. (2014). Digit-Tracking: studying over time the evolution of sensory dimensions of individual product space based on Holistic methods. Sensometrics 2014, Chicago, IL **(The Travel Award of the Sensometric Society)**.

Lê, M.T., Carré, Q., & Lê, S. (2012). A Napping® based methodology for getting an emotion wheel from consumers. EuroSense 2012, Bern, Switzerland.

Lê, M.T., Lê, S., & Nguyen, H.D. (2014). Assessing sensory quality using Holos, an environment for holistic methods. Spise 2014: From senses to quality, Ho Chi Minh City, Vietnam.

Lê, M.T., Carré, Q., & Lê, S. (2012). A Napping® based methodology for quickly getting a model of consumer emotions: A Vietnamese case study. SPISE 2012: Integrating Sensory Evaluation into Product Development: An Asian Perspective, Ho Chi Minh City, Vietnam.

Poster presentations at international conferences

Lê, M.T., & Lê, S. (2014). Holos: a collaborative environment for holistic approaches. Sensometrics 2014, Chicago, IL.

Lê, M.T., & Lê, S. (2014). Holos: a collaborative environment for holistic approaches. Spise 2014: From senses to quality, Ho Chi Minh City, Vietnam.

Lê, M.T., & Lê, S. (2013). A methodology for better understanding consumer emotional states within a cross-cultural context: Where western and eastern cultures meet. Pangborn 2013, Rio de Janeiro, Brazil.

Lê, M.T., Carré, Q., & Lê, S. (2012). A Napping® based methodology for quickly getting a model of consumer emotions. Sensometrics 2012, Rennes, France **(The third best poster of the Sensometric Society)**.

Member of organising committee

Spise 2012 – For further information, please assess the website http://www.conferences.hcmut.edu.vn/spise2012/commitees.html

Spise 2014 – For further information, please assess the website http://www.conferences.hcmut.edu.vn/spise2014/index.php/committees

Appendix 1 Bibliography

A1.1 Studies focus on the comparison of the results obtained by PM and Napping®

Nº	Publication	Product	Sensory method	PM replicates	PM space	Blind duplicates	Descriptors	Statistical analysis	Major findings
1	(Risvik et al., 1994)	5 Commercial chocolates	DA, PM and dissimilarity scaling for pairwise comparison; 1 panel (9 judges): PM training with town's map of Kruskal and Wish (1978)	3	A4 (30 × 21 cm), marked with crossed axes	No	No	GPA RV to compare methods	PM is able to link consumer and sensory studies. High similarity between DA and PM for replicate session 2 and 3 with RV over 0.7.
2	(Risvik et al., 1997)	7 commercial blueberry soups	DA (12 judges) and PM (8 consumers); hedonic rating of consumers after last PM session; PM training with town's map of Kruskal and Wish (1978).	3	A3 (42 × 30 cm)	No	No	PCA on normalized PM data RV to compare methods	High agreement between DA and PM in the first dimension only. Conclude that consumers pull out major differences similar to DA, but not smaller ones. For each PM replicate large individual differences were found.
3	(Barcenas, Pérez Elortondo, & Albisu, 2004)	8 Ewes milk cheeses differing in ripeness level and type	PM (12 consumers) and DA (8 experienced judges); training similar to Risvik et al. (1994), Risvik et al. (1997) using town's map (Kruskal and Wish, 1978); instructed PM panel to use odor, taste and texture attributes.	3	A3 (42 × 30 cm)	No	No	INDSCAL with distances between products and over all replicates. Correlation coefficient between replicate sessions stress and RSQ.	Consumers created very different product maps with little agreement. Lower correlation for consumers than for DA panel.

4	(Perrin et al., 2008)	10 Loire Chenin Blancs differing in vintage and oak ageing	DA (17 judges), PM (12 wine professionals) and FP (12 wine professionals).	No	A2 (60 x 40 cm)	No	From UFP	HMFA on all three methods MFA on PM descriptors as supplementary not categorical variables in analyses.	Similar product maps from all 3 methods
5	(Perrin & Pagès, 2009)	10 Loire red wines from 5 producers and two designations (Anjou Rouge, Anjou Village Brissac)	PM with UFP (14 wine professionals) and DA (8 wine professionals); no formal training for DA.	2	A2 (60 x 40 cm)	No	From UFP	CA and PCA on attributes MFA on coordinates RV to compare methods	Good agreement between maps from UFP and DA Napping procedure pushes judges to find discriminating attributes
6	(Kennedy & Heymann, 2009)	11–12 chocolates with 9 common in all three panels	PM and DA; 3 separate panels (8–9 judges each).	No	60 × 60 cm	No	From UFP	MFA for PM RV to compare methods and panels.	Good correlation between PM and DA. Good correlation between panels. Product separation mostly based on cacao content. Low RV for judges who used a different separation criterion.
7	(Pagès, Cadoret, & Lê, 2010)	8 smoothies – 4 flavours from 2 brands	sorted Napping = PM combined with sorting; 1 panel (24 judges); black glasses and red light.	1	No	No, but considered it	from UFP	HMFA	Expand the PM task with categorisation using the sorting task to group similar samples. Analyse then by HMFA.
8	(Nestrud & Lawless, 2010)	10 Cheddar cheeses, 10 apple varieties	PM and Sorting; 1 untrained panel for each product (19 judges for apples, 21 judges for cheeses); under red light.	No	60 × 60 cm	Yes, 2 for each study	From UFP	MFA RV to compare methods Hierarchical clustering on coordinates from MFA; multiple regression with coordinates of common maps to predict attributes.	PM better than sorting due to better defined clusters in PM. Judges had more difficulties with apples than cheeses.

9	(Dehlholm et al., 2012)	9 Liver pates	Comparison of DA, FP, FMS, PN, GN; 2 professional panels (9 judges each) doing all analyses (only one panel for FP).	No	A2 (60 × 40 cm)	No	From UFP	MFA RV to compare methods Bootstrapping for CI.	Similar maps in all methods DA and PN highest similarity within and between panels.
10	(Louw et al., 2013)	High alcohol (38-43%), divided into 2 sets: Set 1: 5 products + 1R Set 2: 9 products + 1R	Panel 1 (10 subjects): Global Napping (GN) + Partial Napping (PN). Panel 2 (12 subjects): GN & PN. DA (12 subjects), 9products + 2R Note: PN (appearance, aroma, and in-mouth sensations).	Yes, Panel 1: 3 sessions for GN + 3 sessions for PN for each set. No, for Panel 2: 1 session for GN & PN for each set.	A3 (42 × 30 cm)	Yes, 1 product for Set 1, 1 product for Set 2, and 2 products for DA	From UFP	PCA for DA, MFA for GN, HMFA for PN. RV coef. for the comparison of DA, GN and PN.	Product spaces obtained by GN & PN are similar. But they are different from that obtained by DA. Small sample set (n=6), GN & PN are equally reliable, larger sample set (n=10), PN is more stable.
11	(Cruz et al., 2013)	6 probiotic yogurts (4 commercial + 2 prototype products)	CATA, Intensity scales (monadic, balanced sample presentation) Sorting, Napping (a single time) 30 consumers (Comment*: 30 consumers for each task or for all the 4 tasks?).	No	A4 (30 × 21 cm)	No	From UFP	Anova + HCA for intensity scale MFA for Sorting, Napping and CATA MFA, RV coef for comparison 4 methods (based on the coordinates of products on the sensory maps).	Sorting and Intensity scale provide more information of discrimination than Napping and CATA.
12	(Santos et al., 2013)	6 prebiotic mortadellas (meat products),	PM (45 consumers) & UFP (40 consumers).	No	A3 (42 × 30 cm)	No	From UFP	GPA for UFP MFA for PM HMFA for the comparison of 2 methods.	UFP provides a better descriptors (16) than Napping (13).
13	(Ares et al., 2013)	8 orange-flavoured powdered drink samples	Polarised Projective Mapping (PPM, 45 consumers) PM (45 consumers) Polarised Sensory Positioning (triadic	No	A2 (60 x 40 cm)	1	No	MFA	PPM is the combination of PM and PSP. They provided similar product spaces. It means: PPM overcomes the shortcoming of PM as all the samples are presented at the same time.

			PSP, 45 consumers).						Comments: selection of the poles (structure forte).
14	(Kim et al., 2013)	15 green tea commercial products (5 products in each country Korea, China, and Japan)	Napping Consumer acceptability test (15 point scale) Panel 1: 48 Korean subjects, age (20-50). Panel 2: 45 French subjects, age (students).	No	A2 (60 x 40 cm)	No	From UFP	MFA, frequencies of terms are treated as supplementary variable. MFA to compared product space obtained by Korean, and French consumers. Anova, Duncan tests External Pref. Map.	Different sensory perception between Korean and French consumers. Sensory properties were the main factor for consumers, who were more familiar with products. Comment: Knowledge and experience could be different among two groups of consumers (the French consumers as students were recruited).
15	(Cadena et al., 2014)	8 functional yogurts	CATA (19 terms, 81 consumers) PM (81 consumers) Polarized Sensory Positioning (PSP, 3 poles, 81 consumers) DA (they did, but these data were not presented)	No	A3 (42 × 30 cm)	No	From UFP	PCA for DA Cochran's Q test & CA for CATA MFA and bootstrap for PM & PSP.	CATA provides the product space similar to that of DA.; whereas PM provides the least similar product space.
16	(Reinbach, Giacalone, Ribeiro, Bredie, & Frøst, 2014)	8 beers	CATA (38 descriptors, 73 subjects) CATA with intensity (38 descriptors, 63 subjects) Napping (40 subjects) + UFP.	No	A3 (60 x 40 cm)	No	From UFP	A-PLSR, D-PLSR to evaluate the discriminative ability of the methods and descriptors.	Napping is slower and laborious; and suitable for the smaller number of consumers. CATA is faster, less labour-intensive and suitable for lager groups of consumers.

A1.2 Studies focus on the stability of PM and PM and Napping®: statistical methods for analysing the data and interaction task × subject

Nº	Publication	Products	Sensory method	PM replicates	PM space	Blind duplicates	Descriptors	Statistical analysis	Major findings
1	(King et al., 1998)	18 commercial snack bars	Sorting: untrained panel, 24 subjects Structured (labelled axes) & unstructured PM (unlabelled axes): untrained panels, 24 subjects.	No	60 × 60 cm	No	No	GPA and MDS for unstructured PM (Coordinate Averaging) CA on structured PM RV to compare methods.	Coordinate Averaging (CA) less effective than MDS or GPA. Additional information was recovered with unstructured PM and MDA analysis.
2	(Nestrud & Lawless, 2011)	8 artificial symbols (2×2×2 design: 3 factors, 2 levels of each).	PM, 50 subjects	No	60 × 60 cm	No	No	MDS-INDSCAL, MFA.	MDS-INDSCAL and MFA recovered the full dimensionality of the stimuli.
3	(Vidal et al., 2014)	21 data sets of PM (food products) Number of subjects for each data: 26-100.	Napping	//	A3 (42 × 30 cm)	//	//	MFA and confidence ellipses RV coef. Bootstrap	Stability of product space depended on the degree of difference, the type of difference, and the number of samples. 50 consumers is the recommended number for PM study.
4	(Pagès, 2005)	5 Vouvray Chenin Blancs and 5 Touraine Sauvignon Blancs.	DA (8 trained judged) and PM (11 wine professionals).	No	A2 (60 × 40 cm)	//	From DA panel	MFA	10–15% of the judges had problems with the PM task Recommends DA to obtain attributes. Recommends to limit number of products to 12 wines.
5	(Nestrud & Lawless, 2008)	11 Citrus juices from different varieties.	PM; 2 panels (14 chefs and 16 consumers); scaling of attributes; under red light.	No	60 × 60 cm	Yes, 2 juices	Yes, collected after 1st session for scaling exercise in session 2	MFA and GPA on PM coordinates RV to compare methods.	Consumer panel produced similar maps in PM and Scaling, while chefs did not 2 out of 30 judges had difficulties with the PM task.

6	(Kennedy, 2010)	8 Granola bars	PM; 1 untrained panel (15 judges).	3	60 × 60 cm	No	From UFP	MFA, HMFA and PMFA RV to compare results; descriptors as supplementary not categorical variables in analyses.	Little similarity for individual judge between replicate sessions. Speculates if they changed their PM criteria. Overall consensus map was stable over replicate sessions Similar maps with all analytical methods.
7	(Veinand, Godefroy, Adam, & Delarue, 2011)	8 Lemon ice teas from Italy, Switzerland, and France	PM with UFP (30 consumers), FP (43 consumers), RG (42 consumers).	No	A3 (42 × 30 cm)	Yes, one product	From UFP	GPA followed by agglomerative hierarchical clustering on coordinates, RV to compare methods, MANOVA for quality of product separation.	Some consumers had problems with map creation. Suggest to limit PM to expert judges. Difficulties to compare PM with RG and FP due to different numbers of dimensions. Similar product maps from all 3 methods. Consumers were able to detect the duplicated samples.
8	(Torri et al., 2013)	11 red wines	DA (9 trained panel), PM (13 experts), PM (81 consumers) Liking for consumer.	No	A2 (60 × 40 cm)	Yes, 1 product	No	PCA, ANOVA, and plot p*MSE for DA, GPA for PM Internal preference mapping for quality and typical data (expert) and liking data (consumers).	Product space obtained by expert and trained panel are more similar than that of consumers. Liking can be the main criterion for consumers. PM is informative in the condition that assessors have a high level of experience (experts).
9	(Giacalone, Ribeiro, et al., 2013)	9 beers	PN (smell & taste; not appearance & mouth-feel); Panel 1: 8 experts. Panel 2: 9 novices.	No	A3 (42 × 30 cm)	No	From UFP	PMFA: obtains product space, consensus among subjects. HCA for sample classification.	Expert panel has higher agreement than novice panel.
10	(Hopfer & Heymann, 2013)	18 red wines + 2R	Training session PM1 (12 judges - students), 2 groups: square vs. rectangle.	PM1: 3 for each group (3 experiments for each consumer).	PM1: square (70 × 70 cm) and rectangle (81 × 56 cm)	2 (they are different between PM1 & PM2)	DA	CVA & Anova for DA MFA for PM PPI for evaluating individual performance.	Product space depends on the provided space: square vs. rectangle space. People performance index (PPI) to measure the ability of an individual to place blind

			PM2 (22 judges-students): 1 group, 3 shaped spaces.	PM2: No (3 experiments for each consumers).	PM2: square (62 × 62cm), lore (50 × 76cm), shore (76 × 50cm)				duplicated samples close to each other.

(This page is left blank)

Appendix 2 Additional supporting results

A2.1 Permutation-based methods to assess the number of significant dimensions

By definition, a significant dimension is a dimension for which stimuli (i.e., emotions, in this case) are structured in which the positioning of the stimuli is not due to pure chance only. In other words the stimuli are not randomly positioned.

In order to obtain the number of significant dimensions, we compare the eigenvalues of the real data set to those of the permuted data set. The notion of permutation, in mathematics, is related to the act of rearranging objects or values. In this context, the permuted data set corresponds to the situation where subjects have randomly positioned emotions on their rectangles. The idea is to compare the structure of the observed data set to the one of the permuted data sets that is supposed to simulate the absence of structure.

If the eigenvalues of the real data set (obtained from MFA), are significantly higher than the eigenvalues of the permuted data set (obtained from MFA, as well), then we will conclude that emotions are not randomly structured.

Technically, to obtain such permuted data set: (1) for each subject, rows are randomly rearranged, (2) then the couple of coordinates are merged. Such permuted data are simulated a certain number of times (100 times in our analysis) in order to get a distribution of eigenvalues under the hypothesis of random positioning of the stimuli. And finally, eigenvalues from the observed data are compared to the previous distribution of eigenvalues.

A2.2 Representing hierarchy through its sequence of categorical variables, the so-called cognitive process within this context

The coordinates of categorical variable *j* is illustrated on the plane constituted by factor F_s and F_t (cf. Figure A-1). Its coordinate on F_s (resp. F_t) is exactly the correlation ratio between a categorical variable, the variable *j* in this case, and a continuous variable, the variable F_s (resp. F_t) which is the vector of the coordinates of the stimuli of rank *s* (resp. t) (Escofier & Pagès, 2008, p.97-98).

The equation below is the formula to calculate the coordinate of the categorical variable j on the axis F_s of rank s:

$$\eta^2(F_s, j) = \frac{\text{between-clusters inertia}}{\text{total inertia}} = \frac{\sum_{k \in K_j} (I_k/I)(F_s(G_k))^2}{\lambda_s}$$

As mentioned above, F_s denotes the vector of the coordinates of the stimuli of rank s, j denotes a categorical variable. In our context, a categorical variable corresponds to a partition on the stimuli at each hierarchical level. And, G_k denotes barycentre of individuals within modality k (the number of stimuli within a cluster), I_k/I denotes the weight of the categorical variable, and λ_s denotes eigenvalue of rank s. K_j denotes clusters of modalities of variable j.

Values of the correlation ratios lie between 0 and 1.

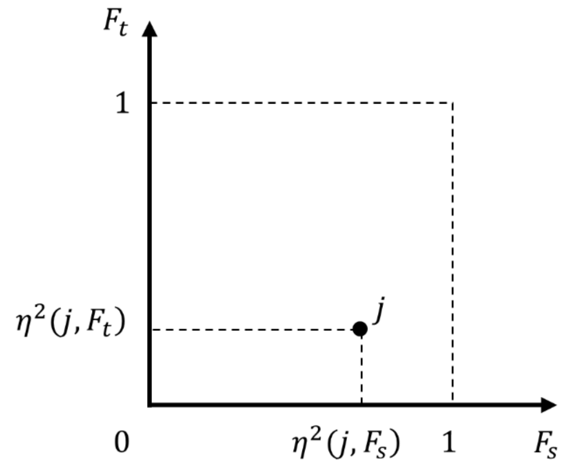


Figure A-1 Illustration of categorical variable on the plane constituted by vectors F_s and F_t by using the correlation ratio.

A2.3 Significant dimensions provided by MFA

Studying inertia of principal dimensions enables us to observe the structure among emotions in one part, and determine the number of "significant" principal dimensions in the other part. In this context, the "significant" principal dimension refers to the eigenvalue explained by the real variation in the real data which is significantly higher than one explained by the random variation in the permuted data.

Figure A-2 below presents the eigenvalues of the fifty-two dimensions obtained from the real data set (black columns) and the permuted data sets (white columns).

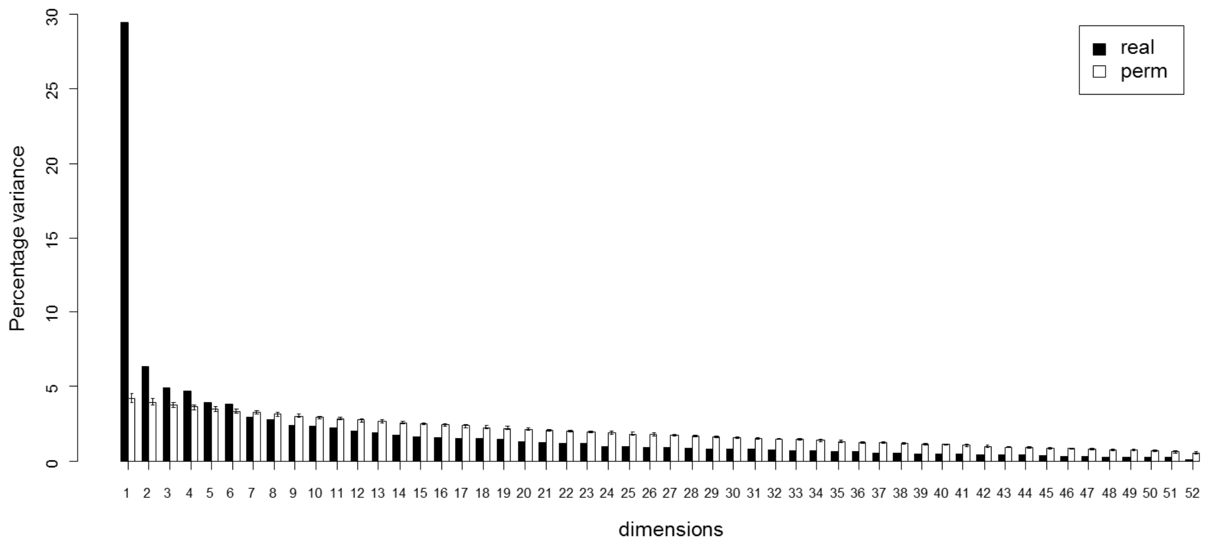


Figure A-2 Eigenvalues of the real data set was illustrated by the black columns and the permuted data sets were illustrated by the white columns obtained by MFA. The eigenvalues of the permuted data sets were illustrated combine with their 95% confidence interval.

The observed data shows the six dimensions that we can take into account when obtaining the emotional space by the fact that they are all significantly different from those of the permuted data. For the emotional space such as that in Figure 2.5, we use the first two dimensions which explain 29.52% and 6.38% of the total inertia.

A2.4 Different strategies for positioning emotions are illustrated on the subject rectangles

Figure A-3 illustrates the four rectangles of the subjects 1, 11, 73, 59. They are chosen because of 56 corresponding to four different strategies for positioning emotions.

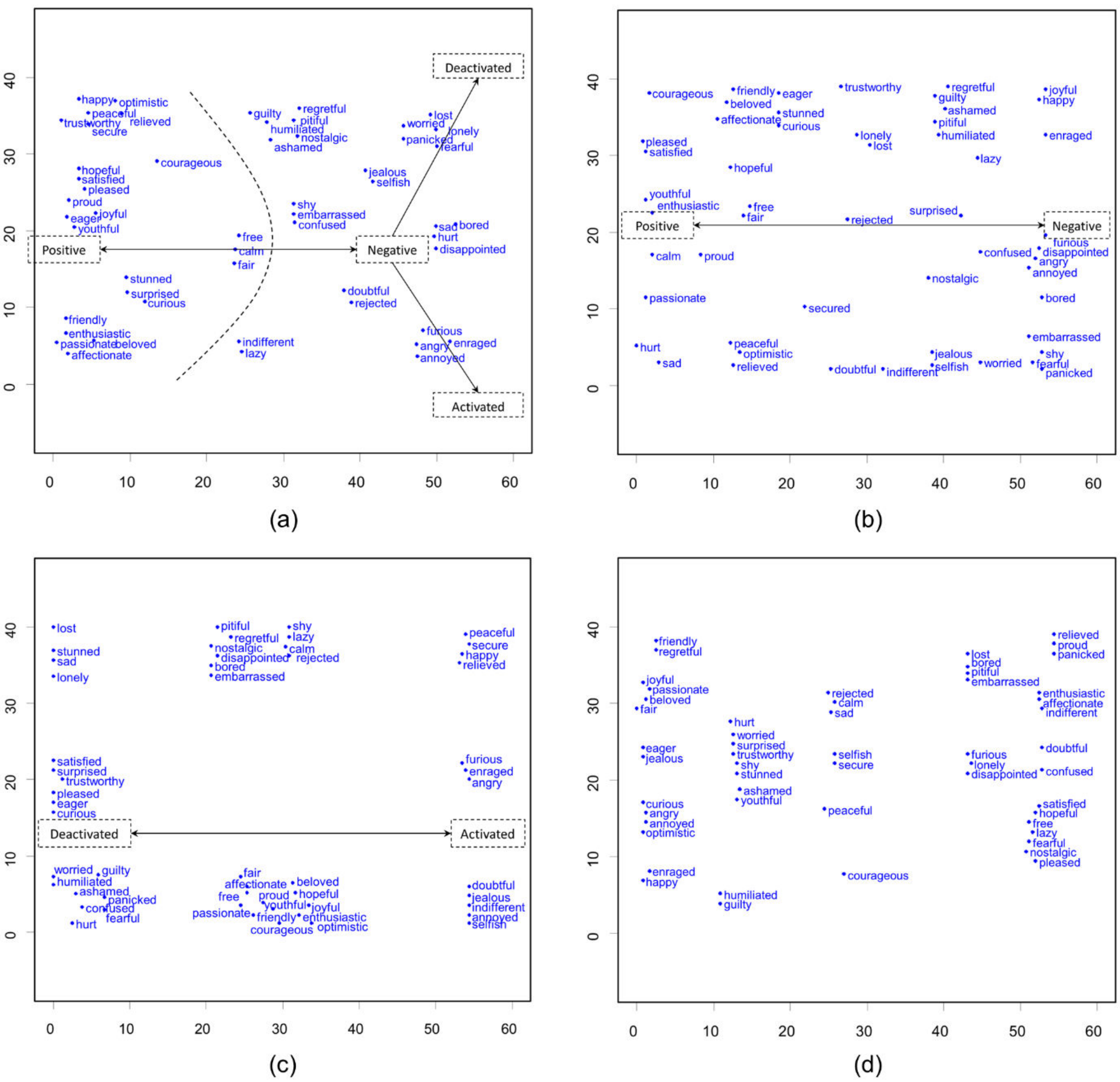


Figure A-3 Illustration of the original rectangles of the subjects 1, 11, 73, 59. (a) rectangle of subject 1, (b) rectangle of subject 11, (c) rectangle of subject 73, and (d) rectangle of subject 59.

(This page is left blank)

Appendix 3 Assessing consumer-perceived product quality by conjoint analysis

This section includes the publication:

Lê, M.T., Lê, S., & Nguyen, H.D. (2014). Assessing consumer-perceived food quality using conjoint analysis. Journal of Science development, Journal of Science and Technology Development, 17-27, Vol.17, ISSN: 1859-0128.

Abstract

Quality is considered the most indispensable means to compete in the marketplace for all food companies. Apart from the objective quality concerning the physical – chemical – microbiological characteristics, which links to the concept of food security; the subjective quality concerning the quality as perceived by consumers is utmost important since it links to the concept of consumer demand (Pagliuca & Scarpato, 2011).

This paper proposed a new methodology which aimed to assess the subjective quality as follows: (1) to examine whether consumers could perceive and/or understand the quality criteria used by experts to assess the product quality, and (2) in case consumers could do, to investigate whether these expert criteria related to the consumer-perceived quality.

This methodology is based on the traditional conjoint analysis combining with an add-on, the sorting task. Consumers are asked to sort 9 pictures of bananas, and then to rate the pictures according to their perception of the perceived quality presenting on the pictures. The methodology will be illustrated through a case study performed on Vietnamese consumers.

The results showed that: (1) consumers could perceive the experts' quality criteria such as bruise, shape, and color; and (2) the consumer-perceived quality related to these expert quality criteria. In addition, the results also showed that bruise was the most important attribute affect consumer appreciation of quality.

For practical application, we expected that this methodology could provide useful information about the subjective quality for those researchers who want to improve the quality based on consumer demand.

Keywords: food quality, sensory defects, holistic approach, conjoint analysis.

Introduction

The basis for all food companies is to improve product quality that insures ongoing consumer appreciation. However, it is not so easy to understand the concept of food quality.

In a more general context, food quality is a multi-faceted concept, which consists of four major aspects, as following: sensory attributes (either positive or defect), health (e.g., nutrition and safety), convenience (e.g., preparation, buying, storing, and eating), and process characteristics (e.g., organic production, animal welfare) (Grunert, 2002). In a simpler context in which solely the intrinsic product attributes (i.e., sensory attributes as in this context) are focused on, the food quality can be considered as overall liking of consumers (Moskowitz, 1995).

Moskowitz & Krieger (1993) published an article entitled "What sensory characteristics drive product quality? An assessment of individual differences". In this, their findings showed that some key sensory attributes could drive the consumer judgment of food quality: in more detail, the taste and flavour could drive the overall liking of fruit pie products. Additionally, these attributes can be used to segment consumers.

Inspired their work, but a bit far from the context of new product design decisions, our starting point aims to determine what sensory defects highly decrease the product quality. More precisely, we are in the context of food quality control, and we want to investigate the relative importance of sensory defects, imputable to each attribute levels, on the overall liking of consumers.

Why is the assessment of food quality from consumers so important? Consumers are key to driving sustainable production and play a central role in sustainable development. This alternative concept a few decades ago has become the mainstream today.

Rather than adopting the work of Moskowitz & Krieger (1993) in the field of quality control, we propose in this article a new methodology for assessing the notion of quality perceived by consumers. This methodology is based on the traditional conjoint analysis combining with an add-on, the sorting task.

This minor add-on seems to be important in our context since we want to ensure that whether consumers can perceive the sensory attributes (i.e., the defects) used by the experts in the one part, and whether they can use these attributes to structure the overall liking in the other part.

This methodology consists of three main steps: constructing different product profiles from the sensory defects by means of experimental design (1), assessing the perception (2), and assessing the consumer-perceived quality (3). This methodology will be demonstrated through a "simple" case study on the quality of banana performed by Vietnamese consumers.

Methodology

The methodology consisted of three main steps: (1) choosing the expert criteria in order to formulate products, (2) exploring structure of product space perceived by consumers by using sorting task, and (3) investigating the relationship between the notion of quality perceived by consumers and the expert criteria by using quality rating task. This methodology was demonstrated through a "simple" case study on evaluating the quality of banana performed by 30 Vietnamese consumers.

Along with the main steps of the methodology, we also introduced the Holos environment for collecting data for the sorting and quality rating experiments, as well as the way for analysing such data.

Three main steps of the methodology

Step 1: constructing the product concepts using experimental design

An advantage of experimental design enables to investigate the structure of the profiles more systematically. By varying these profiles by an experimental design, then presenting them to subjects in a known way, researchers can understand which factors work and which do not work. Moreover, researchers can also estimate the relative contributions of these factors in the product space.

In our case study, the four factors (i.e., the expert criteria used to evaluate the visual quality of banana), such as colour, shape, bruise, and cigar-end rot were chosen from the technical guidelines proposed by Dadzie & Orchard (1997). Each factor has three levels.

For varying the profiles, a full factorial design, which includes all combinations of the factor levels, could be the first-thinking solution. However, this design was not practical since the total number of combinations is large: with 34 equals 81 combinations (i.e., the profiles) that could be constructed. This high number of profiles might lead subjects to be unduly burdened when they are asked to provide judgments on all profiles. To overcome this short-comings, using fractional factorial designs, which can reduce the number of products to be administered to subjects, are recommended.

Here, we constructed the profiles following a fractional factorial design – the Graeco Latin square (Lorenzen & Anderson, 1993). Following this design, there are 9 profiles (equals 3^{4-2}) which could be constructed from 4 factors, each at 3 levels (cf. Table A3 - 1). These profiles are combinations of levels of the factors for which they are designed as illustrated in Figure A3 - 1.

Step 2: assessing consumer perception using sorting task

Most of previous conjoint applications ignore this step because the concepts of their factors and levels seem to be clearly perceived by consumers. In other words, these factors and levels could be understood without trouble, for instance: average waiting time (10, 20, or 30 minutes) or price for a bus ticket (1, 1.5, or 2€ per hour), etc. Nevertheless, in this research context, we remark that consumers deal with the expert criteria. Thus, a natural question arises whether or not consumers can perceive the expert criteria. To do so, the sorting task was used.

The sorting aims to investigate how consumers perceive intrinsically the conceptual similarity between the profiles. In this experiment, the null hypothesis is: consumers cannot perceive the expert criteria. If this was the case, the product space could not be interpreted by the expert criteria. In other words, consumers might not understand the sensory defects that will be structured in the next steps. Vice versa, the alternative hypothesis is: consumers can employ the expert criteria. If this was the case, the product space could be interpreted through the expert criteria. In other words, consumers might understand the sensory defects that will be structured in the next steps.

Table A3 - 1: A 3^{4-2} Graeco Latin square design for 4 factors, each at 3 levels.

Product	Colour	Shape	Bruise	Cigar-end rot
1	1	1	1	1
2	1	2	3	2
3	1	3	2	3
4	2	1	3	3
5	2	2	2	1
6	2	3	1	2
7	3	1	2	2
8	3	2	1	3
9	3	3	3	1
1	green	high heterogeneity	high bruise	high cigar
2	green-yellow	heterogeneity	bruise	cigar
3	yellow	homogeneity	non-bruise	non-cigar

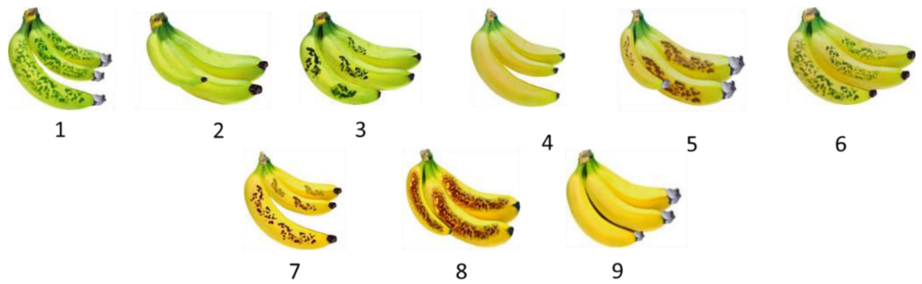


Figure A3 - 1 Products generated based on the 3^{4-2} Graeco Latin square design

The sorting was performed in two sessions. In the first session, each consumer was asked to sort 9 profiles into as many groups as he/she would like to, assuring that the products in the same group were perceived as the same, and the products in different groups were perceived as different. In order to get information about the sensory attributes of each group, a verbalization task was required for each consumer in the second session.

Based on the final configurations (i.e., the partitions) obtained by subjects, a data table, whose rows represent the profiles, whose columns represent the subjects, and whose intersections represent the group that a profile belongs to, was obtained.

Among the possible statistical techniques for analysing the sorting task data, e.g. MDS or DISTATIS, FAST (Cadoret, Lê, & Pagès, 2009) was chosen. The main advantage of FAST is that: it takes into account individual partition in the one part, and takes the same importance for each consumer when structuring the product space in the other part. Technically, FAST is a combination of MCA (which provides the consensual representation of the products, i.e. the product space) and MFA (which provides the representation of the subjects).

Step 3: assessing the notion of quality using conjoint analysis

Once the consumers could perceive the expert criteria (as shown in section 3.1 of this paper), we move to the next question: whether or not consumers use the expert criteria to structure the product quality. To answer this question, the traditional conjoint analysis was performed to investigate the relative importance of sensory defects on their notion of quality.

Different from the traditional conjoint analysis (Rao, 2014), in this experiment, all the profiles were presented simultaneously to subjects, not in a serial monadic fashion. The reason for that is because: from our point of view, once a subject evaluates all profiles simultaneously, his/her personal criteria corresponding the quality may not differ from one profile to the others. In practice, the profiles are randomly presented to subjects and they are asked to arrange these profiles on a continuous scale.

In terms of analysing the data, the attribute trade-offs will be estimated following the part-worth model (cf. Eq.1):

$$Y = \beta_0 + \sum_{i,j} \beta_{ij} X_{ij} + \varepsilon \qquad \text{(Eq.1)}$$

Where i denotes the number of attributes (i=1, 2, 3, 4), j denotes the number levels per attributes minus 1 (j=1, 2 – each attribute has three levels), and β denotes the unknown parameters (i.e., the part worth functions).

All the part worth functions are estimated by using ordinary least squares (OLS) regression method: the quality score was considered dependent variable, the levels of sensory defects were considered dummy variables. The interactions, in this model, were confounded with the error term.

A case study and implemented software

The proposed methodology will be demonstrated through a case study on the quality of banana, performed by 32 Vietnamese consumers. They are all students, whose age ranges from 18 to 24, at the Ho Chi Minh City University of Technology, Vietnam.

All the subjects are asked to complete the sorting task first, then the quality rating task. In terms of collecting data, both the sorting and quality rating task were performed by using Holos environment (Lê & Lê, 2014). Subjects conducted experiment using a tablet device: stimuli were displayed as icons that could be dragged with a finger. Further information about Holos could be found on http://napping.agrocampus-ouest.fr/

All statistical analyses were performed using R software, version 3.1.2 (R Core Team, 2014) with the package FactoMineR (François Husson et al., 2014) for analysing the sorting data, and the package conjoint (Bak & Bartlomowicz, 2012) for performing the traditional conjoint analysis.

Results

Did consumers understand the expert criteria?

Figure A3 - 2 shows the representation of 9 product profiles obtained from MCA. This representation was the consensus over all consumers on the first factorial plane constituted by the first two dimensions of variability.

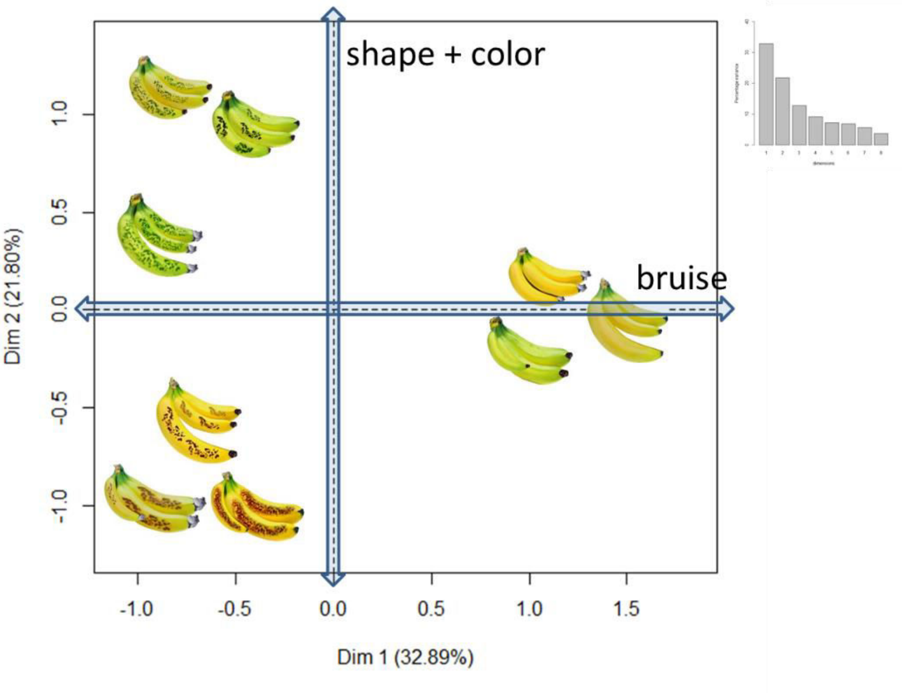


Figure A3 - 2 Consensual representation of product space perceived by consumers

The first dimension, denoted Dim.1, explained 32.89% of total variability, whereas the second dimension, denoted Dim.2, explained 21.8% of total variability. Although the results of verbalization task were not presented in this paper, Figure 2 showed that Dim.1 opposed the products bruise on the left side to the products non-bruise on the right side of the plane, and Dim.2 opposed the products heterogeneity (levels 1&2 of the shape factor) with yellow colour on the bottom left side to the products homogeneity with green colour on the top left side of the plane.

These two dimensions explained a high percentage of the total variability, which indicated that there was a consensus among subjects. Interestingly, the interpretation of the two dimensions corresponded to the criteria bruise, shape and colour used by experts in evaluating the quality of banana. In other words, we can conclude that consumers could perceive, and in fact they used, the expert criteria to sort the products.

How did consumers structure the relative importance?

Table A3- 2 shows the part worth functions, i.e. the parameters of the part worth model obtained by OLS. The R-squared explains 49.22% of the variance in the dependent variable (i.e., quality score), which probably be considered good fitness of the part worth model. The results show that the part worth functions of the high and medium bruise levels, the heterogeneity level (of the shape), and the green colour level are significant different from zero.

What these results imply is that the most important defect makes banana less attractive perceived by consumers is the bruise: its high bruise level will decrease 173.1 points, and its medium bruise level will decrease 71.2 points over total of 1000 points of the quality of banana. The second important defect is the shape: its high heterogeneity level will slightly decrease 2.6 points of the quality, but its medium heterogeneity will decrease 106.3 points of the quality. Then, the third important defect is the colour: its green level will decrease 43.5 points of the total quality.

Finally, although the effect of the cigar-end rot defect is not significant, we can say that its high cigar level decreases 16.5 points of the total quality.

Table A3- 2 The part worth functions obtained by OLS

```
Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)         451.594     12.267  36.813  < 2e-16 ***
factor(x$color)1    -43.500     17.349  -2.507   0.0127 *
factor(x$color)2     16.688     17.349   0.962   0.3369
factor(x$shape)1     -2.656     17.349  -0.153   0.8784
factor(x$shape)2   -106.385     17.349  -6.132 2.95e-09 ***
factor(x$bruise)1 -173.125     17.349  -9.979  < 2e-16 ***
factor(x$bruise)2  -71.240     17.349  -4.106 5.29e-05 ***
factor(x$cigar)1   -16.500     17.349  -0.951   0.3424
factor(x$cigar)2    28.146     17.349   1.622   0.1059
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 208.2 on 279 degrees of freedom
Multiple R-squared:  0.4922,  Adjusted R-squared:  0.4776
F-statistic:  33.8 on 8 and 279 DF,  p-value: < 2.2e-16
```

It is important to note that the two dummy variables were used to define for the three-level attribute. The dummy variables, for the third level, are automatically drop, i.e., their values are zero. The reason for that is because, from the statistical point of view, if all the dummy variables were kept, this would give a model with non-invertible $X'X$ matrix where X is the design (or model) matrix. Consequently, we could not estimate the parameters of the model because the sum of all the dummy variables is the constant predictor, which equals 1.

However, the part worth model can be written with the intercept and all dummy variables in the case that we impose a restriction on the model, the so-called effects coding. The restriction here is: the sum of all dummy parameters is scaled to zero. This restriction might be useful if we need to rescale some the coefficients (within one factor), to make them comparable.

Table A3- 3 shows all the part worth functions when using effects coding (their values are rounded to one decimal place). Remarkably, the third level of attributes focuses on the attractive of the quality, for instance, if the banana is at the non-bruise level, the quality will increase 244.3 points; or it is at the homogeneity level, or at the yellow colour level, the quality will increase 109 or 26.8 points, respectively. The part worth functions for each attribute can be graphically demonstrated through a bar graph (cf. Appendix).

Table A3- 3 The part worth functions of all levels

	Levels	P-W function.values
1	intercept	451.5
2	green	-43.5
3	green-yellow	16.6
4	yellow	26.8
5	high.heterogeneity	-2.6
6	heterogeneity	-106.3
7	homogeneity	109.0
8	high.bruise	-173.1
9	bruise	-71.2
10	non.bruise	244.3
11	high.cigar	-16.5
12	med.cigar	28.1
13	non.cigar	-11.6

For a better interpretation, this fitted model can be rewritten with all the part worth functions as follows:

$$Y = 451.5 - 43.5 \times \text{green} + 16.6 \times \text{green.yellow} + 26.8 \times \text{yellow} - 2.6 \times \text{high.heterogeneity} - 106.3 \times \text{heterogeneity} + 109.0 \times \text{homogeneity} - 173.1 \times \text{high.bruise} - 71.2 \times \text{med.bruise} + 244.3 \times \text{bruise} - 16.5 \times \text{high.cigar} + 28 \times \text{med.cigar} - 11.6 \times \text{low.cigar}$$

Figure A3 - 3 shows the relative importance of attributes in a total utility function. This relative measure was obtained by rescaling these measures such that they all add to 100%. This result shows that bruise is the most important sensory defect, which contributes 45.62% of total quality. Then, the second important defect is the shape, which contributes 23.45% of total quality. The colour and the cigar-end rot contribute respectively 17.68% and 13.25% of total quality.

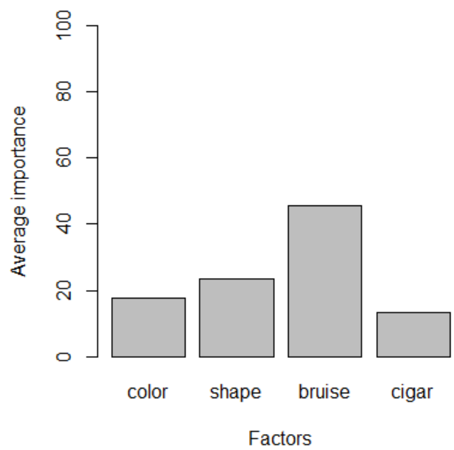


Figure A3 - 3 Relative importance of sensory defects on the quality of banana

Conclusion

This paper shows an evidence that the two most common beliefs about consumers and consumer-perceived quality, at least in the context of quality control, are wrong:

- Consumers cannot perceived the quality criteria used by experts.
- Consumer-perceived quality does not have structure and it does not related to the expert criteria.

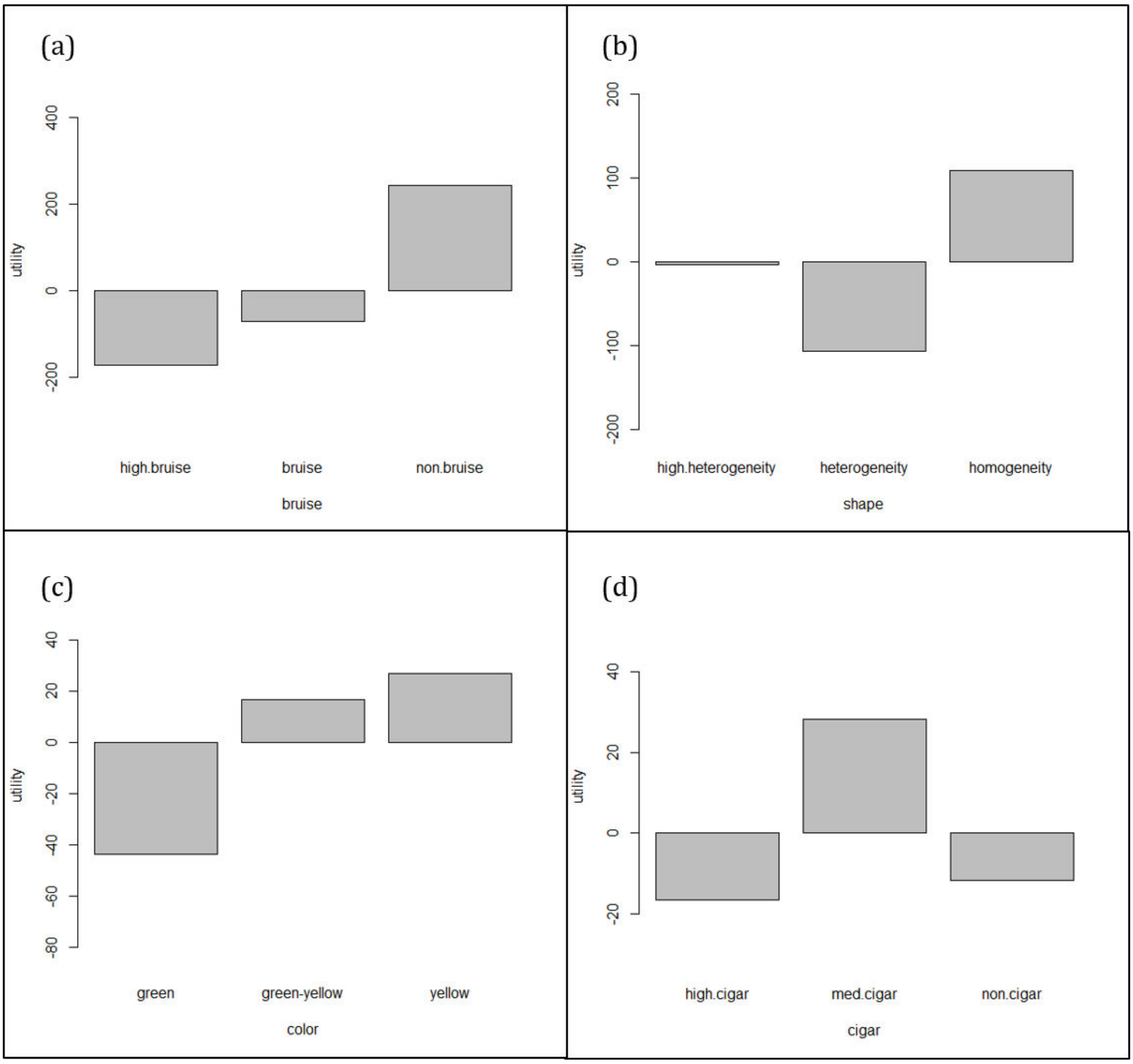
Besides, the proposed methodology provides valuable information about key sensory attributes driven consumer-perceived quality in the one part; and about the part worth model, which aids in interpreting the relative contribution of each attribute as well as its levels affect the quality in the other part.

In terms of the data collection, the two experiments of this methodology were performed by using Holos environment. By observing subjects and doing surveys, subjects showed great interest in using tablet: it is intuitive, easy, and entertaining.

We expected that this methodology could be a useful tool for practitioners and researchers who want to study the consumer-perceived quality in order to improve the total quality of product.

Appendix: The part worth function for the quality of banana:

(a) bruise, (b) shape, (c) color, and (d) cigar



(This page is left blank)

Appendix 4 Holos: A collaborative environment for similarity-based holistic approach

This section includes the publication:

Lê, M.T., Brard, M., & Lê, S. (2016, accepted). Holos: A collaborative environment for similarity-based holistic approaches, Behaviour Research Methods.

Abstract

This article aims to introduce Holos – a new collaborative environment that allows researchers to carry out experiments based on similarity assessments between stimuli, such as projective mapping and sorting tasks. An important feature of Holos is its capacity to assess the real-time individual process during the task.

Within Holos environment, researchers can design experiments on its platform, which can handle four kinds of stimuli, such as: concepts, images, sounds, and videos. In addition, researchers can share their study resources within the scientific community, such as: stimuli, experimental protocol, and/or data collected. With the dedicated Android application combined with a tactile human machine interface, subjects can conduct experiments using a tablet to obtain similarity measures between stimuli. On the tablet, stimuli are displayed as icons that can be dragged with one finger to position them depending on the way they are perceived. By recording the *x*-,*y*-coordinates of the stimuli while subjects move the icons, the obtained data can reveal the cognitive processes of the subjects during the experiment. Such data, named digit-tracking data, can be analysed with the SensoMineR package.

In this article, we describe how researchers can design an experiment, how subjects can perform an experiment, and how digit-tracking data can be statistically analyzed within Holos environment. At the end of the article, a short exemplar experiment is presented.

Keywords: Holos; Android application; Digit-tracking; Similarity assessment

Introduction

To investigate how people perceive stimuli in terms of similarities, holistic approaches such as projective mapping or sorting tasks can be used. For the projective mapping task, subjects are asked to position stimuli on a plane according to their similarities (Risvik, McEwan, Colwill, Rogers & Lyon, 1994). The closer two stimuli are positioned, the more similar they are perceived. In terms of data collected, each subject provides a dissimilarity matrix between the stimuli based on a quantitative measure, namely the distance between the stimuli. This technique is also known as Spatial Arrangement Procedure (SAP; Goldstone, 1994). For the sorting task, subjects are asked to gather stimuli into clusters according to their similarities. Two stimuli belong to a same group if they are perceived as similar. In terms of data collected, each subject provides a dissimilarity matrix between the stimuli based on a binary qualitative measure, namely 0 if two stimuli belong to a same group, and 1 otherwise.

A common feature between these two methods is that subjects assess a set of stimuli based on their perceived similarities, hence the concept of similarity-based methods (Valentin, Chollet, Lelièvre, & Abdi, 2012). These two methods can be considered as holistic in the sense that each stimulus is considered as a whole and not as a sum of predefined components. These methods are widely used in different fields such as sensory analysis and marketing, where stimuli are globally assessed through human senses.

The objectives of such methods are to provide a representation of the stimuli based on a group of subjects, and eventually to compare the individual assessments within that representation.

In terms of the data collection, projective mapping and sorting tasks can be easily performed on various types of stimuli, due to commercial and non-commercial softwares. For instance, EyeQuestion (Logic8, 2015), Fizz version 2.5.0, (Biosystème, 2015), and TimeSens (ChemoSens, 2015) is designed to handle stimuli in form of simple texts and images; TCL-LabX (Gaillard, 2009) – a free software product designed to handle sound stimuli only, but has to be programed for texts or images; and NappingPlayer (Robitza, Helpiquet, Kitzinger & Hlavacs, 2013) – a free video player for Android tablets designed to conduct Napping® (Pagès, 2005) for videos only. With the current software solutions, researchers must adapt their choice to the type of stimuli.

Nevertheless, two main drawbacks subsist with the existing softwares. First, they can only handle some, not all stimuli in the form of texts, images, sounds, or videos. Consequently, this limits the application of these methods to various research fields. Second, only final configurations provided by subjects are collected. Consequently, they do not allow the recording of the different steps followed by the subjects to create their final configuration: indeed, a given dissimilarity matrix provided by two subjects can reflect two different behaviors during the task (shown later in this article). This sequence of steps, which can be assimilated to the cognitive process used by the subject to perform the task, can bring a new perspective to the collected data.

In this article, we describe a tool to collect the individual process that leads a subject to their dissimilarity matrix and we describe an R package to illustrate this individual process. This tool, named Holos, is the combination of a *software product* for collecting data and *a server* for storing data. Compared to most commercial softwares, Holos can handle several kinds of stimuli: texts/concepts (i.e., words or statements), images, sounds, and videos. In this, the *software product* is an Android application with a tactile human machine interface in which subjects can conduct experiments using a tablet. During the experiment, as the subject moves an icon by dragging it across the tablet screen, the server records the trajectory of the subject's finger (i.e., the trajectory of the stimulus) in the form of digit-tracking data, akin to eye-tracking data. The digit-tracking data can be analyzed with a function available in the SensoMineR package (Lê & Husson, 2008). Interestingly, the server is a free collaborative environment, in which researchers with a Holos account, can share partially or totally their study resources within the scientific community. In the next sections, we will present in details Holos and its different features; we will end the article by presenting an exemplar experiment.

Holos

Holos can be defined as a software solution that allows researchers to measure the real-time individual process during an experiment based on holistic similarity assessments between stimuli. Holos was designed to handle four types of stimuli: concepts, images, sounds and videos.

Overall structure

Figure 1 represents the working process of Holos environment, which can be broken down into three components. First, a server in which researchers can design their experiment. On this server, researchers can also share their experiment resources (stimuli, protocol, etc.). In that sense, Holos can be defined as a collaborative environment dedicated to holistic data issued from methods based on the measurements of between-stimuli similarities or differences. Second, an Android application combined with a tactile interface with which the participants can perform the experiment. Holos records the streaming *x*-,*y*-coordinates of the stimuli when they are dragged by the participant on the tactile tablet during the task. This feature, which is the most important characteristic of Holos, generates data that allow researchers to review their experiments and then to better understand the behavior of their subjects during the task. These data have been called digit-tracking data, as Holos has been developed for tactile tablets, in other words as stimuli are dragged by the subject's digit when performing the task. Finally, the third component is a function available in the SensoMineR package to analyze such data. When the holistic task consists in collecting quantitative similarities-dissimilarities data (i.e., when the task is a projective mapping task or one of its variants), the digit-tracking data can be statistically analyzed in order to obtain a graphical representation that could be assimilated to the cognitive process of a subject during the task.

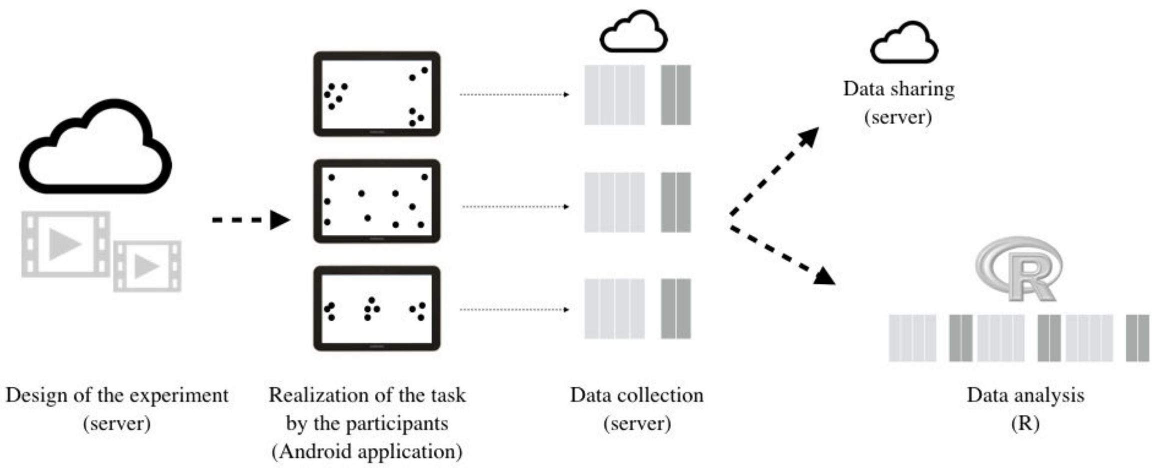


Figure A4-1. Representing steps in the working process of Holos environment.

Installation

Each component of Holos can be considered independently. First, researchers can access the server at the platform http://napping.agrocampus-ouest.fr by creating an account. Second, participants can download the Android application on their tablet device. This application, named *NSubject*, is also available at the website http://napping.agrocampus-ouest.fr. Its installation is similar to that of other applications in the Google Play store. Third, to analyze digit-tracking data, researchers can use the SensoMineR package (Lê & Husson, 2008), a free R package dedicated to the analysis of sensory data.

Designing an experiment

Experiments are designed by the researchers via their account on the website http://napping.agrocampus-ouest.fr. Practically, in terms of the experiment organization, researchers can manage their stimuli and protocols, and the recruitment of their subjects for their experiment. A user manual is available on the website.

Specifying the stimuli – As mentioned previously, Holos can handle four formats of stimuli (viz., text, image, audio and video) as Holos is dedicated to the study of concepts, images, sounds and videos. For a given experiment, stimuli must be the same type/format. When stimuli are images, as the stimuli will be represented as small icons during the experiment on the tablet device, the researcher has to import the stimuli and their thumbnails as well. Once stimuli are uploaded, the researcher can change the names that will appear to the subject during the experiment, directly below the thumbnails. By default, it is the name of the stimuli that have been imported. They can be replaced for instance by random three digits numbers or by letters.

Specifying the protocol – The protocol corresponds to the instructions that will pop-up in a splash screen on the tablets at the beginning of the experiment. As Holos can handle several kinds of tasks (projective mappings, sorting tasks, etc.), the protocol should be specified carefully.

Specifying the participants – As a web application, the researcher has the possibility to send an email to a list of subjects to invite them to join a given experiment, and to perform the experiment whenever they want, on their own tablet. To do so, the researcher has to enter the list of subjects to be contacted. Once this list is filled an email is sent to the subjects to inform them that they are invited to an experiment. The invitation is associated with the ID number of the experiment, which is mandatory for the subject to perform the experiment.

Performing an experiment as a participant

To perform an experiment, the subject has to open the *NSubject* Android application. The subject has to provide their name, email address and the ID of the experiment. Once the application is opened, the protocol of the experiment appears in a splash screen, and the stimuli are represented in the form of icons. The icons are randomly displayed on the left side of the screen. This arrangement is different from one subject to the other. Each icon can be dragged by using one's finger. A double-click on an icon induces an action that depends on the type of stimuli: it makes the image appear in case of image stimuli, plays the sound or the video in case of audio and video stimuli, or opens a screen in which appears the definition of the concept in case of text stimuli. When the task is over, subjects can write down information to describe the stimuli.

Digit-tracking data and their analysis

During the task, as the subject is moving the stimuli on the tablet, their coordinates are recorded. More precisely, for a given subject, the application is generating a matrix of dimensions $I \times 2$ each time a stimulus has been moved, where I is the number of stimuli and 2 the number of coordinates recorded (i.e., the x-axis and the y-axis). By recording all these coordinates, it is possible to reconstitute the trajectories of all the stimuli. These data are automatically stored in the Holos server.

Once the experiment is over (i.e., once all the participants have performed the task), the researcher can access the digit-tracking data in the server. To analyze them, the folder obtained during the experiment must be downloaded on the researcher's computer. Digit-tracking data can be analyzed using the SensoMineR package (version available at the website http://sensominer.free.fr/). The *holos* function of this package allows to analyze such data. As the data depends on the holistic method used to collect them, the main argument of this function is the type of task that was performed. When the holistic task is the data collection based on quantitative similarities-dissimilarities (i.e., when the task is a projective mapping or one of its variants), representations of the cognitive processes can be obtained: this is not the case when the task is a sorting task or one of its variants. The analysis of the data obtained with a projective mapping task or one of its variants is described below.

The *holos* function analyzes the final configurations. In the matrix containing the final configurations, each line corresponds to a stimulus and the columns correspond to the final configurations provided by the subjects (i.e., the x-,y-coordinates provided by each subject at the end of the experiment) (cf. Table 1). Such a matrix is analyzed with a Multiple Factor Analysis where each group is composed of the two columns relative to a subject. This analysis provides two major pieces of information: the representation of the stimuli over all subjects, and second, the representation of the subjects.

Table A4-1 Dataset containing the final configurations.

	X_{fin} - S1	Y_{fin} - S1	X_{fin} - S2	Y_{fin} - S2	X_{fin} - S3	Y_{fin} - S3	X_{fin} - S4	Y_{fin} - S4	X_{fin} - S5	Y_{fin} - S5
Sti. *A*										
…										
Sti. *i*										
…										
Sti. *P*										

The *holos* function analyzes the digit-tracking data. In the matrix containing the digit-tracking, each line corresponds to a stimulus and the columns correspond to the *x*-,*y*-coordinates of the stimuli during the steps of realization of the experiment by the subjects (cf. Table 2). Such a matrix is analyzed with a Multiple Factor Analysis where each group is composed of the two columns relative to one step followed by a subject. In this analysis, only the final configurations are defined as active groups whereas all the intermediate steps are defined as supplementary groups. This analysis provides, for each subject, a representation that allows to understand how this subject has perceived the stimuli throughout the experiment. This representation can be assimilated to the mental processing used by this subject to perform the task until the getting of his final configuration. The *holos* function of SensoMineR provides other pieces of information to summarize the digit-tracking data for each subject (e.g., number of steps, duration of the task, and duration of the learning phase).

Table A4-2 Dataset containing the stimuli trajectories

	*X*0–*S*1	*Y*0–*S*1	*X*1–*S*1	*Y*1–*S*1	…	*X*fin–*S*1	*Y*fin–*S*1	…	*X*fin–*S*5	*Y*fin–*S*5
Sti. *A*										
…										
Sti. *i*										
…										
Sti. *P*										

Experiment

This experiment was realized with stimuli corresponding to cards created with a factorial design. The aim of such an experiment was: (1) to understand which factors were first perceived by subjects, and (2) to understand how the subjects perceived the stimuli during the whole experiment (i.e., to determine the mental processing the subjects used during the task to provide their final configuration).

Method

Stimuli – 16 cards were created with a factorial design. The following were the factors used: background color (orange vs. yellow), shape color (blue vs. green), contour (dotted *vs.* continuous line), shape (circle *vs.* square), size (big *vs.* small), position (top *vs.* bottom), and background pattern (shaded *vs.* plain). Figure 2 presents two of the 16 stimuli.

Protocol – Participants performed a Napping® task, a particular case of projective mapping. The protocol was the following one: "During this experiment, you are asked to position the stimuli on the screen in such a way that two stimuli are close if you perceived them as similar and two stimuli are distant if you perceived them as different.”

Participants – We will use the data from five subjects.

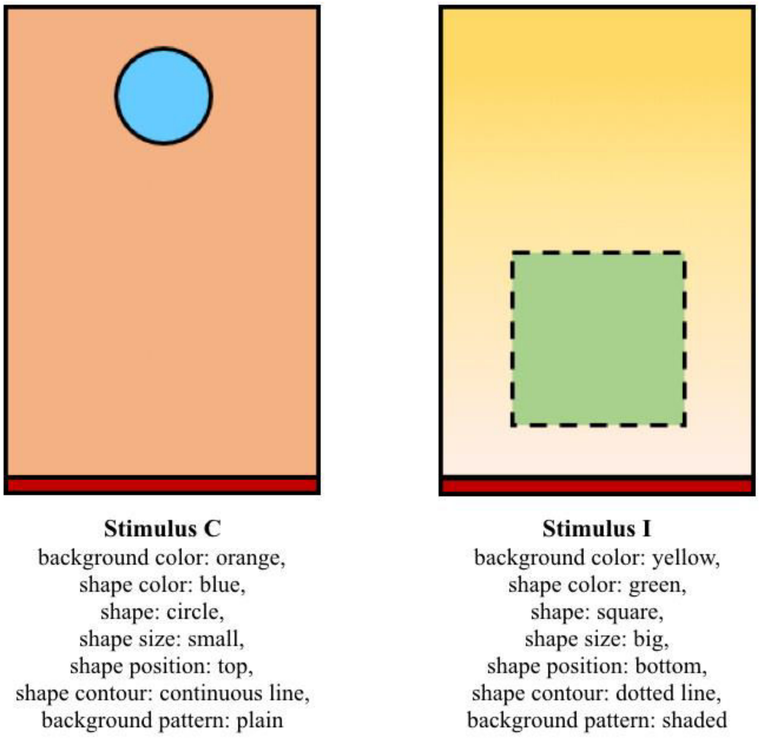


Figure A4-2. Illustration of two stimuli C and I created with a factorial design

Results

Final configurations – The five subjects produced different final configurations (cf. Figure 3). Subject S1 separated the stimuli based on their background color; subjects S2 and S5 used the shape color; subject S3 used both background and shape color; only subject S4 used the contour to separate the stimuli.

As shown in Figure 4a, the first dimension of variability among the cards is the shape color and the second dimension is the background color. The first dimension, denoted Dim 1, opposes the cards with green shape on the left side with blue shape on the right side. It explains 43.23% of total variability of the data. The second dimension, denoted Dim 2, opposes the cards with orange background on top to the cards with yellow background at the bottom. It explains 36.19% of total variability of the data.

The representation of the subjects involved in the experiment in Figure 4b has to be interpreted jointly with the previous representation of the stimuli. For instance, S1 gets a high value on Dim 2, meaning that the main variability on their final configuration is consistent with the background color. In other words, S1 separated the stimuli based on the background color. S2 and S5 get a high value on Dim 1 since they separated the stimuli based on the shape color. S3 gets a high value on both Dim 2 and Dim 1 since they separated the stimuli according to both background and shape color. Finally, S4 gets low values on both Dim 1 and Dim 2 since he separated the stimuli based on the contour.

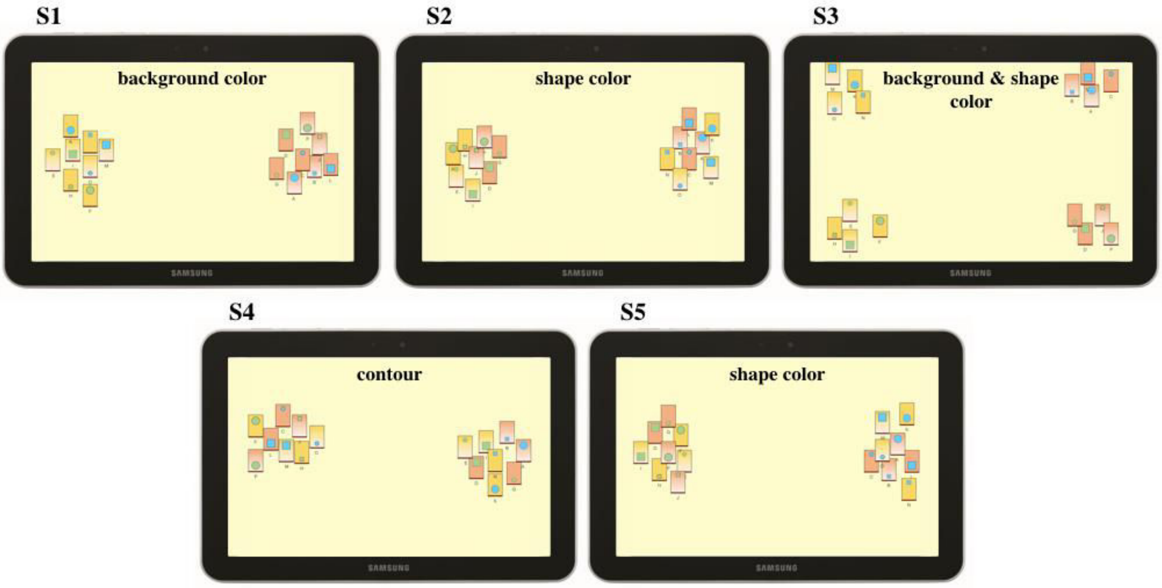


Figure A4-3. Final configurations provided by the five subjects

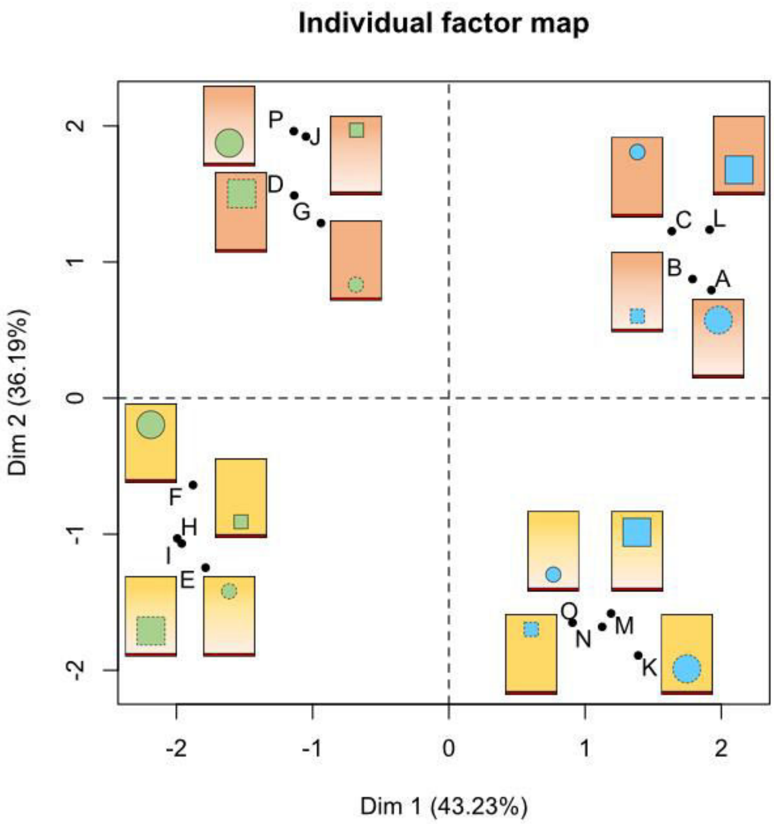


Figure A4-4a. Representation of the stimuli

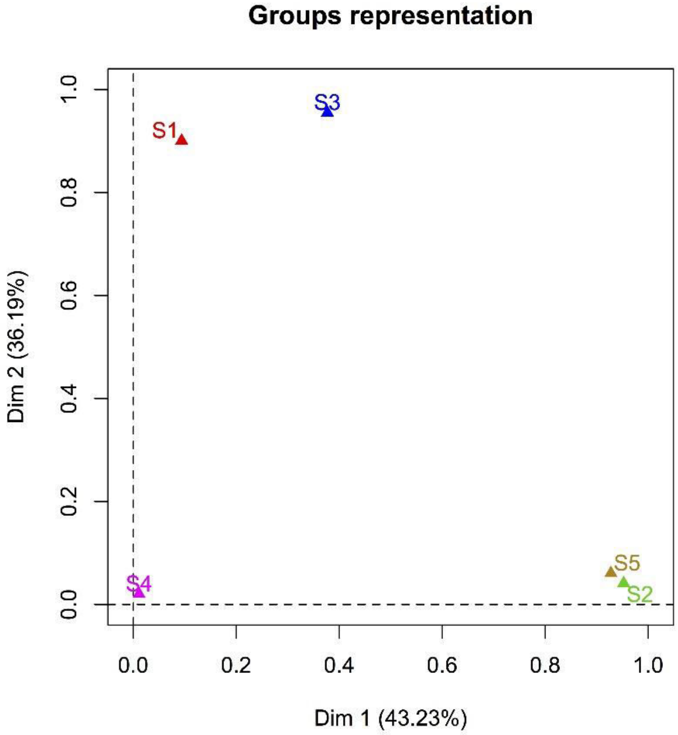


Figure A4-4b. Representation of the subjects

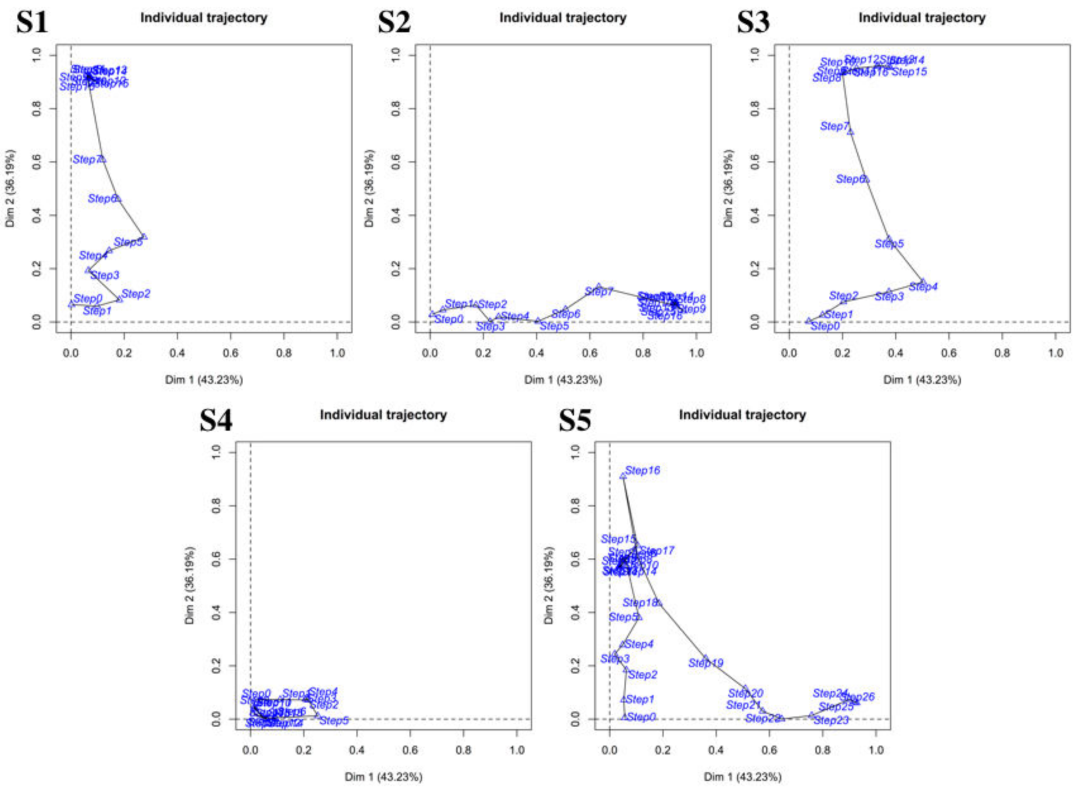


Figure A4-5. Representations of the cognitive processes

Cognitive processes – The representations of the cognitive processes of the subjects involved in the cards experiment are shown in Figure 5. The representation of the cognitive process used by S1 shows that this subject performed 16 steps to obtain their final configuration. At the beginning of the experiment, the configurations of the cards (*e.g*. Step 1 or Step 2) get low values on Dim 2. Gradually, at the end of the experiment, the configurations of the cards (*e.g.* Step 15 or Step 16) get high values on Dim 2. In other words, the cognitive process used by this subject shows that they gradually perceived the background color in successive steps from the beginning to the end of the experiment.

The representation of the cognitive process used by S3 shows that this subject took 16 steps to reach their final configuration. This subject perceived both the shape color and the background color throughout the experiment to separate the stimuli.

The representation of the cognitive process used by S4 shows that this subject conducted 16 steps to obtain their final configuration but they neither perceived the background color nor the shape color during the whole experiment.

The representations of the cognitive processes used by S2 and S5 is interesting. As mentioned previously, these two subjects are very close in the representation of the subjects (cf. Figure 4b), meaning that the way they have positioned the stimuli is similar. Even if they have provided a same final configuration, S2 and S5 did not have the same cognitive process when performing the task. First, S2 and S5 did not use the same number of steps to reach their final configuration: it took 16 steps to S2 and 26 steps to S5. Second, S2 and S5 did not use the same mental processing when they performed the task. Indeed, S2 gradually perceived the shape color in successive steps from the beginning to the end of the experiment whereas S5 first perceived the background color (from Step 1 to Step 16) and then perceived the shape color (from Step 17 to Step 26).

Discussion on experiment

This example highlights the importance of collecting digit-tracking data during an experiment based on similarity assessments between stimuli. Indeed, the sole analysis of the final configurations would have been misleading. Without the representation of the cognitive processes, we would have concluded that subjects S2 and S5 performed the task similarly and saw the same similarities and dissimilarities within the stimuli. This conclusion would have been incorrect as subjects S2 and S5 behaved differently during their task: they did not use the same number of steps to reach their final configuration, nor did they use the same mental processing.

Conclusion

In the present article, we presented Holos, an environment for holistic experiments based on similarity assessments between stimuli. We exposed the general features of this environment and we described how the data it provides can reveal real-time individual process during the task. By recording the *x*-,*y*-coordinates of the stimuli, Holos is the first software based on similarity measures that integrates the subjects' behavior into the data analysis step. The importance of recording these stimuli trajectories, called digit-tracking data, has been highlighted with an exemplar experiment: digit-tracking data can notably reveal how two subjects can provide identical final configurations with two different cognitive processes.

Beyond the fact that Holos is a software product for collecting real-time individual process, Holos is also a collaborative and a free environment. On the platform, researchers can access open study resources (stimuli, protocols, etc.) and benefit this feature of Holos to conduct cross-cultural studies.

The environment still needs to be improved, in particular when researchers want to share their study resources. In terms of stimuli, when an experiment involves videos, subjects may encounter problems especially when the size of the video is large and when a subject wants to play it back. In terms of methodologies, we want to add new ways of collecting data and in particular the Q-sort, which allows to study subject subjectivity when they are confronted to a pool of stimuli.

Since the beginning of its development, dozens of experiments have been performed successfully within the Holos environment. Our hope is that Holos will be used by researchers from various fields such as experimental psychology, cognitive science, marketing, and so forth.

(This page is left blank)

Appendix 5 Posters presented at conferences

A5.1 A Napping® based methodology for quickly getting a model of consumer emotions

A napping based methodology for quickly getting a model of consumer emotions

[1]Minh-Tâm Lê, [1]Quentin Carré, [1]Sébastien Lê
[1]AGROCAMPUS OUEST, Laboratoire de Mathématiques Appliquées, Rennes
65 rue de Saint-Brieuc, CS 84215, 35042 Rennes Cedex

Introduction and context: Why do emotions matter in consumer research?

Nowadays, the Holy Grail that companies are pursuing is a better understanding of the consumer experience. This concept encompasses every aspect of evoked emotions that a consumer has with a product and a supplier over the duration of usage (Gentile, Spiller & Noci, 2007). In other words, it notably reflects the relation between a consumer and a product from an emotional point of view.

Although emotions have gained increasingly attention in psychology and consumer research, there's still room for new methodologies that handle the particular case where consumers are confronted to a large range of emotions they have to assess. This presentation proposes a methodology based on the Napping® (Pagès, 2005) to get a model that represents how emotions or emotional states are structured and organized between them.

Methodology

Napping®, a projective test *par excellence*

"In psychology, a projective test is a type of personality test in which the individual offers responses to ambiguous scenes, words or images."

In order to reveal how emotions are perceived unconsciously by consumers we use the Napping® method such as a projective test.

The emotions: how large is large?

As the adage goes "To Choose is To Renounce". One main feature of our methodology lies in the fact that we present to the consumers an exhaustible selection of every possible emotion in daily life.

Remarks:

1. To handle such a large range of emotions we perform the Napping® on a giant screen.
2. This "non selection" will allow us for instance to compare cultural differences.

A Vietnamese case study

"The proof is in the pudding". The methodology was experimented on 120 Vietnamese consumers and a selection of 53 emotions.

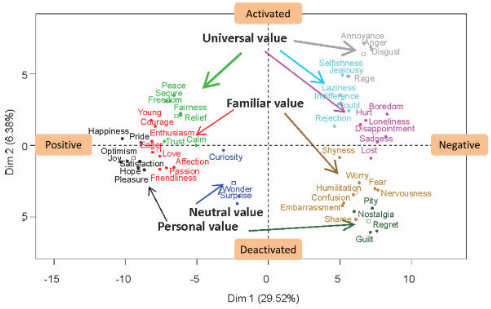


Fig. 1: The consensual representation of the emotions

An emotional model in the form of a factorial plan

The consensual representation of the emotions (cf. Fig. 1) provided by the Napping® shows:

1. An opposition between positive and negative emotions.
2. An opposition between activated and deactivated emotions, as commonly seen in the literature (Posner, 2005).
3. A typology of nine groups in accordance with the values and the characteristics of the Vietnamese culture (Phan, 1998, pp.32-107).

The emotions wheel

G1: positive emotions of personal value.
G2: positive emotions of familiar and social value.
G3: universal value of positive states.
G4: neutral value
G5: universal value of defect.
G6: universal value of non-isolation.
G7: universal value of non-forceful reaction.
G8: negative emotions of familiar and social value
G9: negative emotions of personal value.

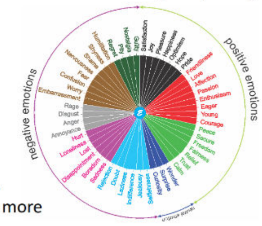


Fig. 2: The emotions wheel

We introduce a new representation to visualize the typology of emotions in the form of a wheel. This wheel has a very singular property, as two sectors associated to emotional states are all the more close than emotional states were felt as similar for the consumers.

Based on the experiment, we may conclude that the use of the Napping® on a large range of emotions is feasible for consumers. It is then possible to get a consensual representations of the emotions as well as a wheel of emotions.

Comparing cultural differences can be done by applying the proposed methodology combined with HMFA.

ACKNOWLEDGEMENT The authors would like to thank François Husson (Agrocampus Ouest) for the comments that help improve the research.

References

Gentile, C., Spiller, N., Noci, G. (2007). How to sustain the customer experience: An overview of Experience Components that co-create value with the customer. *European Management Journal*, **25** (5), 395–410.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances Using multiple factor analysis : An application to the study Of 10 white wines from the Loire Valley, *Food Quality and Preference, 16*, pp. 642-649.

Posner, J., Russell, J.A., Peterson, B.S. (2005). The circumplex model of affect: an tergartive approach to affective neuroscience, cognitive development, and psychopathogy, *Dev. Psycholopathol, 17*, pp. 715-734.

Phan, N. (1998). Bản sắc văn hóa Việt Nam, NXB. Văn hóa thông tin, 578 trang.
(Phan, N. (1998). Characteristics of Vietnamese culture, Culture Information Publishing House, 578 pp.)

Please cite as: T.M. Lê, Q. Carré, S. Lê (2012). **A Napping® based methodology for quickly getting a model of consumer emotions**. Poster presented at Sensometrics 2012, Rennes, France.

A5.2 A methodology for better understanding consumer emotional states within a cross-cultural context: Where western and eastern cultures meet

A methodology to better understand consumer emotions within a cross-cultural context: where western and eastern cultures meet

AGRO CAMPUS OUEST · ueb UNIVERSITÉ EUROPÉENNE DE BRETAGNE

The aim of this research is to further explore the work of Lê *et al.* (2012) in which they proposed a methodology to obtain an empirical model of the way emotions are structured. In this study, this methodology is applied and extended to a cross-cultural comparison between two seemingly opposite cultures: the western culture, represented by France, and the eastern culture, represented by Vietnam.

In a preliminary step, we directly apply the methodology proposed by Lê *et al.* (2012) to each culture in order to obtain a graphical representation of the emotions issued from the dimensional approach, and a graphical representation of the cognitive process that leads to the formation representation of the emotions issued from the categorical approach. Then, we compare the cultures with respect to those two previous representations, thanks to a Hierarchical Multiple Factor Analysis (HMFA) (Le Dien & Pagès, 2003).

Methodology

Data collection: 100 French subjects and 120 Vietnamese subjects performed a Napping® task in which they were asked to position 53 emotions on a 60 cm by 40 cm rectangle according to the way they perceived similarities among the emotions.

Data analysis:

- From the study of Lê *et al.* (2012), the representation from the dimensional approach is obtained by using Multiple Factor Analysis (MFA) (Pagès, 2004); the representation from the categorical approach is attained by using Divisive Algorithm (MacNaughton-Smith *et al.*, 1964) where clusters are divided into two sub-clusters iteratively (from the whole set to the singletons). This algorithm is applied on the results issued from MFA.
- For the cross-cultural comparison, a HMFA is performed on the data collected from 220 subjects. In this HMFA, at the first level, subjects are balanced within the country they belong to; at the second level, countries are balanced.
- Data were analyzed using the R software version 2.15.3. Multiple Factor Analysis and Diana Algorithm for Hierarchical Divisive Clustering were performed by the FactoMineR package version 1.23 (Husson *et al.*, 2013) and the cluster package version 1.14.3 (Maechler *et al.*, 2012), respectively.

Results

Dimensional approach for Vietnam and France | **Categorical approach for Vietnam and France**

Methodology of Lê *et al.* (2012)

In Fig.1, for both countries, the first dimension opposes the positive emotions to the negative emotions. The second dimension opposes, among the negative emotions, the activated emotions to the deactivated emotions.

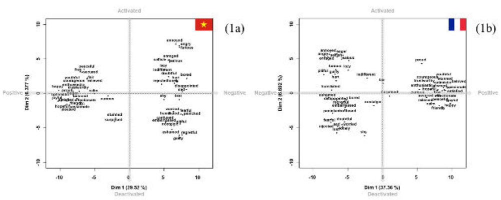


Fig.1a (*resp.* Fig.1b): The representation of the emotions from the 120 Vietnamese (*resp.* 100 French) subjects, obtained from the dimensional approach. The two first dimensions explain respectively 29.52% and 6.38% (*resp.* 37.36% and 6.69%) of the total variance.

The two representations, in Fig.2a and 2b, are directly derived from the representations of the dimensional approach in Fig.1 (*cf.* Lê *et al.*, 2012). They are interpreted by starting from the top of the tree to its bottom.

Both countries seem to have separated first the positive emotions from the negative emotions; then among the negative emotions, the activated emotions from the deactivated emotions.

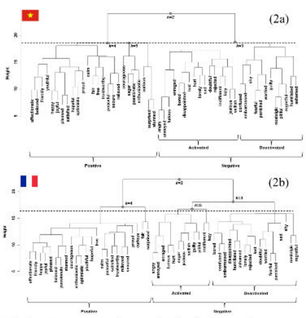


Fig.2: Representations obtained from the categorical approach by (a) the Vietnamese subjects and (b) the French subjects.

Hierarchical Multiple Factor Analysis (HMFA)

The structure of the emotions provided by the Vietnamese subjects is similar to the one provided by the French subjects, except for the two emotions *eager* and *stunned* (Fig.3).

	VN	FR
Eager	Positive	Negative
Stunned	Neutral	Positive

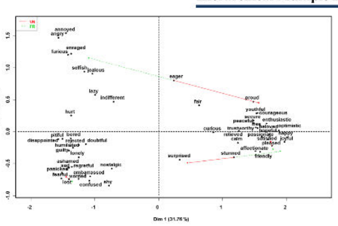


Fig.3: Partial representations of the emotions obtained by HMFA.

The hierarchy, as in Fig.2a and 2b, is associated with the cognitive process in which we consider each step of the hierarchy as a partition on the stimuli.

In order to quickly compare such cognitive process, we project its sequence of steps on the axes issued from HMFA (Fig.4).

A comparison of the two cognitive processes reveals that the obtained clusters are akin between the two cultures, at least for the first two important steps of the hierarchy.

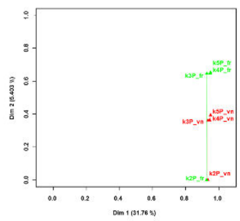


Fig.4: Cognitive processes of the Vietnamese subjects (P_vn) and the French subjects (P_fr). k2 corresponds to a classification in 2 groups, k3 corresponds to a classification in 3 groups…

Conclusions

This study extended the findings of the previous study concerning the empirical model of emotions (Lê *et al.*, 2012) in a cross-cultural context. In details, this study provides an additional understanding of the structure of emotions for a given culture in the one hand, and the way for comparing such structures among cultures in the other hand. Understanding the sole emotions, culture by culture, must be the starting point for anyone who wants to integrate emotions in product development given a cross-cultural context.

Authors: Minh-Tâm LÊ & Sébastien LÊ
Address: Agrocampus Ouest, 65 rue de Saint-Brieuc, 35042 Rennes Cedex, France
Email: minh-tam.le@agrocampus-ouest.fr

References

Husson, F., Josse, J., Lê, S., & Mazet, J. (2012). FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R. R package version 1.20. Available from http://CRAN.R-project.org/package=FactoMineR

Lê, M.T., Carré, Q., & Lê, S. (2012). A napping based methodology for getting an emotion wheel from consumers. Abstract book of The 5th European Conference on Sensory and Consumer Research (Oral presentation).

Le Dien, S. & Pagès, J. (2003). Hierarchical Multiple factor analysis: application to the comparison of sensory profiles, Food Quality and Preferences, 18, 453-464.

MacNaughton-Smith, P, W.T. Williams, M.B. Dale, & L.G. Mockett. (1964). Dissimilarity analysis: A new technique of hierarchical subdivision, Nature, 202, 1034-1035.

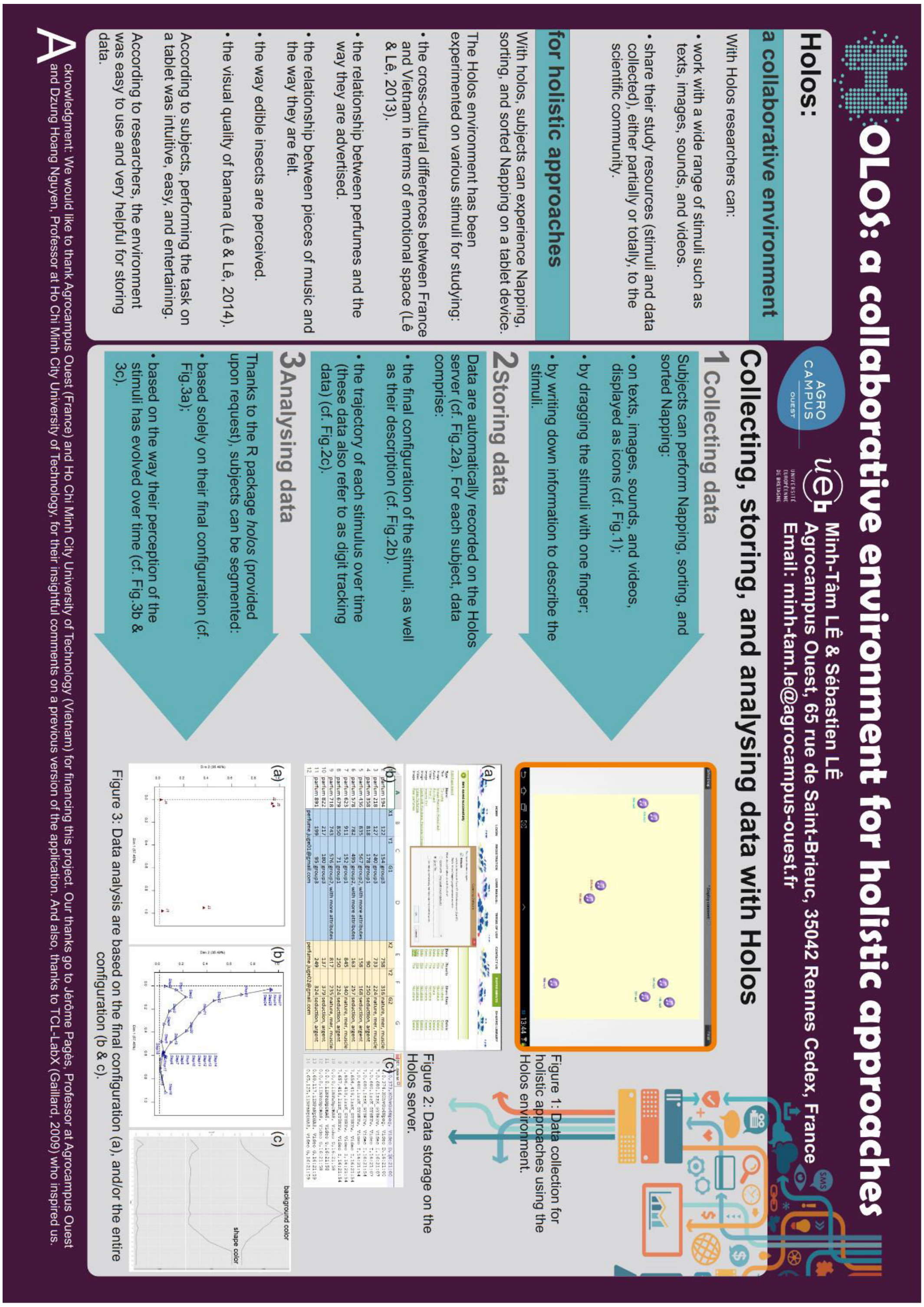
Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2012). cluster: Cluster Analysis Basics and Extensions. R package version 1.14.3.

Pagès, J. (2004). Multiple factor analysis: Main features and application to sensory data. Revista Colombiana de Estadistica, 27, 1-26.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire Valley, Food Quality and Preferences, 16, 642-649.

Please cite as: T.M. Lê & S. Lê (2013). **A methodology for better understanding consumer emotional states within a cross-cultural context: Where western and eastern cultures meet**. Poster presented at the 10th Pangborn, Rio de Janeiro, Brazil.

A5.3 Holos: a collaborative environment for holistic approaches



Please cite as: Lê, M.T. & Lê, S. (2014). **Holos: a collaborative environment for holistic approaches**. Poster presented at Sensometrics 2014, Chicago, IL, USA.