

Formation d'ingénieur en science des données

Année 2020 / 2021

Valence, compréhension, adéquation
des termes sensoriels

Rapport bibliographique présenté par :

Laudine Bertrand

En partenariat avec la Société Française d'Analyse Sensorielle



Maître de stage :

Arnaud Thomas (SensoStat)

GT méthodes :

Ronan Symoneaux (ESA)

Léa Koenig (ESA)

Odile Parizel (Bonduelle)

Massimo Cealti (Altitude C)

Table des matières

I. Valence	1
a) Définition	1
b) Méthodes manuelles de mesure de la valence d'un mot.....	1
c) Méthodes automatiques de mesure de la valence d'un mot.....	4
d) Logiciels de mesure de la valence d'un texte.....	7
e) Mesure des émotions et produits.....	8
II. Compréhension	11
a) Définition	11
b) Compréhension et terminologie sensorielle de texture.....	11
c) Compréhension et allégations de santé.....	14
d) Compréhension et descripteurs du vin.....	17
e) Compréhension et terminologie médicale.....	18
f) Autres études de la compréhension.....	20
g) Comparaison entre experts sensoriels et consommateurs.....	22
III. Adéquation	22
a) Définition	22
b) Adéquation des descripteurs sensoriels.....	22
c) Alternative : tâche de catégorisation.....	24
Bibliographie	

I. Valence

a) Définition

« La valence correspond à la valeur hédonique de l'expérience, à savoir son caractère plaisant ou déplaisant » (Blanc, 2006).

La valence se caractérise comme une réponse émotionnelle, positive ou négative, à une situation ou à un stimulus pouvant appartenir à toutes les modalités sensorielles. Sa dimension se définit sur un continuum plaisir/déplaisir, opposant les états agréables (e.g. joie, satisfaction, contemplation) aux états désagréables (e.g. malheur, contrariété, désespoir) (Mauss & Robinson, 2009 ; Bradley & Land, 2000).

L'analyse de la valence peut concerner une multitude d'éléments tels qu'un corpus de texte, un paragraphe, une phrase, ou encore un mot. Elle peut être mesurée à l'aide de plusieurs méthodes, regroupées en deux catégories : les méthodes manuelles et les méthodes automatiques.

b) Méthodes manuelles de mesure de la valence d'un mot

La majorité des mesures de la valence se base sur des instruments d'auto-évaluations, qui permettent à l'individu d'exprimer ce qu'il ressent. Ces méthodes sont plus adaptées au domaine de l'évaluation sensorielle en raison de la familiarité des consommateurs avec ces systèmes de notation et de leur facilité de mise en œuvre.

❖ Les méthodes verbales

Plusieurs travaux ont permis l'évaluation des caractéristiques émotionnelles des mots, tels que ceux de Vikis-Freibergs (1976), Messina et al. (1989), Syssau et Font (2003) ou encore Painchaud (2005). Dans la grande majorité des évaluations de la valence, les juges donnent leurs réponses sur des échelles ordinales bipolaires en plusieurs points ; la valence s'exprimant ainsi sur un continuum représenté par des nombres naturels (e.g. 1 = très négatif, 5 = très positif) ou des nombres relatifs (e.g. -5 = très négatif, 5 = très positif).

- **Vikis-Freibergs (1976)** a sélectionné 398 mots pour lesquels 100 participants ont répondu à la tâche d'évaluation émotionnelle à l'aide d'un système de notation en trois temps. Il a d'abord été demandé aux participants de décider si le mot était émotif ou neutre, puis, si le mot était jugé émotif, de déterminer s'il était positif ou négatif et de le coter sur un continuum en quatre points « peu », « assez », « très », « extrême ». En plus d'être la première norme émotionnelle

pour des mots en langue française, ce travail a permis de démontrer la constance des composantes émotionnelles au cours du temps ; les évaluations effectuées par Vikis-Freibergs étant significativement corrélées avec les plus récentes.

- **Messina et al.** (1989) ont fait appel à 209 participants afin d'évaluer 904 substantifs de la langue française à l'aide d'une échelle allant de 1 à 5 (1 = « désagréable, antipathique, déplaisant, connotation négative », 5 = « agréable, sympathique, plaisant, connotation positive »). Les 904 mots, divisés en 3 questionnaires, ont été évalués par les juges dans un délai de 24 heures. La particularité de cette étude repose dans la répartition des participants en trois groupes d'âges (17 – 29 ans, 30 – 59 ans, 60 – 90 ans) et dans la répétition de certains mots dans les questionnaires. L'homogénéité transgénérationnelle de l'évaluation de la valence ainsi que la constance dans l'évaluation individuelle de la connotation des mots ont ainsi pu être démontrées.
- **Syssau et Font** (2003) ont procédé à l'évaluation de la valence de 604 mots, dont 348 mots concrets (e.g. banc, arbre) et 256 mots abstraits (e.g. douleur, joie). Cette dernière a été réalisée par 600 adultes, âgés de 18 à 26 ans. Trois listes ont été formées avec les 604 mots, comportant, à parts sensiblement égales, des mots concrets et abstraits. Chaque mot a été jugé par au moins 100 participants. Deux évaluations différentes ont été réalisées ; une première évaluation sur une échelle nominale en trois points (négatif, neutre, positif) et une deuxième sur une échelle ordinale bipolaire en onze points, allant de -5 (très négatif) à 5 (très positif). Grâce à cette étude, Syssau et Font ont préconisé l'utilisation d'une échelle en onze points en raison de la précision apportée, principalement pour la saisie de la nuance des mots neutres. Elles ont également démontré que les mots neutres étaient plutôt concrets, et que les mots émotifs étaient plutôt abstraits.
- **Painchaud** (2005) a établi un corpus de 2401 mots et demandé à des étudiants d'évaluer une fiche de 30 mots. L'évaluation émotionnelle des mots n'a pas été réalisée à l'aide d'une échelle ordinale bipolaire classique, mais de deux colonnes positives (« + », « + + ») et de deux colonnes négatives (« - », « - - »). Notons qu'il était précisé dans les consignes d'indiquer l'affectivité suggérée par chaque mot de manière intuitive, sans y réfléchir longuement.

Grâce au coefficient de corrélation linéaire de Bravais-Pearson, les normes présentées ci-dessus ont été comparées entre elles. Pour la valence, les corrélations entre ces dernières sont toutes supérieures à 0,87 ($p < 0,01$) et sont présentées dans le Tableau 1 (Leveau et al., 2011).

	<i>Messina et al (1989)</i>	<i>Syssau et Font (2003)</i>	<i>Painchaud (2005)</i>
<i>Vikis-Freibergs (1976)</i>	0,91	0,90	0,88
<i>Messina et al. (1989)</i>	-	0,94	0,92
<i>Syssau et Font (2003)</i>	-	-	0,92
<i>Painchaud (2005)</i>	-	-	-

Tableau 1 : Corrélation entre les valences des normes présentées (Leveau et al., 2011)

La forte corrélation entre les résultats a permis la constitution d'une métanorme de 5656 entrées évaluées selon la dimension de la valence. Cette métanorme est utilisée par l'outil EMOVA (Leveau et al., 2011) pour l'évaluation émotionnelle de textes, se basant sur l'hypothèse que la valence d'un texte peut être déterminée à partir de la valence de chaque mot qui le constitue.

❖ Les méthodes graphiques

A la différence des méthodes décrites ci-dessus, les questionnaires avec images émotionnelles permettent aux individus de décrire leurs émotions à l'aide d'images et non de textes.

- **Bradley et Lang (1999)** ont procédé à l'évaluation de près de 1000 mots de la langue anglaise, portant sur la valence, l'activation physiologique et le contrôle exercé par le participant sur le sentiment ressenti. Pour cette étude, chaque participant disposait de 56 mots à analyser grâce à une échelle non-verbale des dimensions émotionnelles appelée SAM (Self-Assessment Manikin). Cette dernière illustre, pour chaque dimension (valence, excitation, dominance), neuf niveaux d'évaluation grâce à un personnage graphique dont l'état est décliné en neuf degrés (Figure 1). Pour la valence, les extrémités sont caractérisées par un visage souriant et enjoué pour l'une, et par un visage morose et grimaçant pour l'autre (Kaneko et al., 2018).

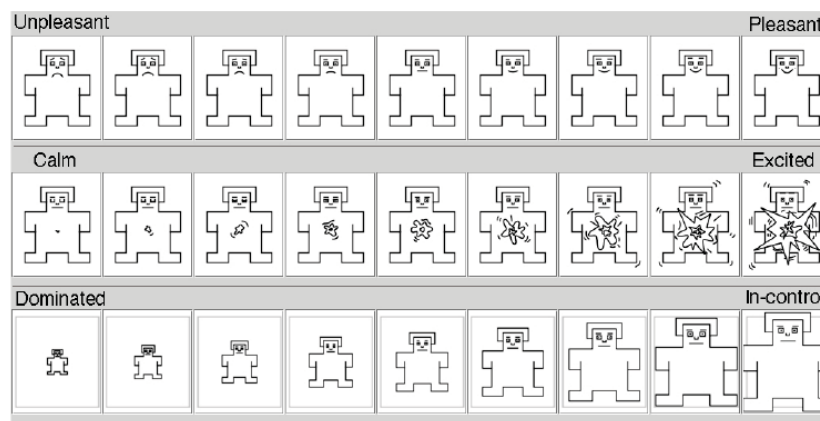


Figure 1 : Echelles SAM (Self-Assessment Manikin)

Afin d'exploiter les résultats obtenus avec cette étude, les 1000 mots anglais évalués ont été traduits en français grâce au dictionnaire WordReference, puis retraduits du français vers l'anglais à l'aide du dictionnaire Collins. Ainsi, les évaluations des 563 mots, ayant été retraduits de façon identique, ont été comparées avec celles réalisées par différents auteurs francophones. Concernant la valence, les évaluations des termes français sont significativement corrélées avec celles des termes anglais après double traduction (Tableau 2) (Leveau et al., 2011).

	<i>Vikis-Freibergs</i> (1976)	<i>Messina.et al</i> (1989)	<i>Syssau et Font</i> (2003)	<i>Painchaud</i> (2005)
<i>Bradley et Lang (1999)</i>	0,81	0,87	0,92	0,93

Tableau 2 : Corrélation entre les valences des normes françaises et celle de la norme anglaise après double traduction (Leveau et al., 2011)

Selon une étude portugaise réalisée par Soares et al. (2012), les femmes ont tendance à donner des notes plus extrêmes que les hommes pour la valence, c'est-à-dire qu'elles évaluent les mots désagréables comme étant plus négatifs et les mots agréables comme plus positifs. Cela suggère que la réactivité aux stimuli émotionnels est plus élevée chez les femmes que chez les hommes (Monnier & Syssau, 2013).

Outre les méthodes détaillées ci-dessus, il existe également d'autres méthodes, pour lesquelles les réponses des participants sont involontaires, incontrôlables ou inconscientes. Les études du système nerveux autonome, des réactions de sursaut, des activations des hémisphères cérébraux ou encore de la posture corporelle peuvent être citées comme exemples pour ces méthodes de mesure de la valence (Mauss & Robinson, 2009).

c) Méthodes automatiques de mesure de la valence d'un mot

Dès 2002, Turney, Littman et Bestgen ont proposé des méthodes automatiques d'évaluation de la valence de n'importe quel terme (adjectif, nom, verbe) grâce à sa proximité sémantique avec d'autres mots dont la valence est connue (Bestgen, 2006).

- **Turney et Littman** (2002, 2003) ont développé la procédure **PR-ASL**, qui analyse la proximité sémantique entre le mot à analyser et 14 mots germes dont la valence est connue (7 à valence négative - mauvais, méchant, médiocre, négatif, malheureux, faux, inférieur - et

7 à valence positive - bon, gentil, excellent, positif, heureux, correct, supérieur -). Cette méthode n'utilise qu'un dictionnaire de valence de 14 mots et l'étend en déterminant la proximité sémantique entre ces mots et n'importe quel autre mot. Avec cette méthode, un mot est d'autant plus négatif qu'il est proche des points de repère négatifs et éloigné des points de repère positifs (Bestgen, 2006 ; Vincze & Bestgen, 2011).

La proximité sémantique est mesurée grâce à l'Analyse Sémantique Latente (ASL), « une technique mathématique qui vise à extraire un espace sémantique de très grande dimension à partir de l'analyse statistique de l'ensemble des cooccurrences dans un corpus de textes. Le point de départ de l'analyse est un tableau lexical qui contient le nombre d'occurrences de chaque mot dans chaque segment de textes. Ce tableau fait l'objet d'une décomposition en valeurs singulières qui en extrait les dimensions orthogonales les plus importantes » (Bestgen, 2006).

La similarité sémantique entre le mot cible et le germe est mesurée grâce au calcul du cosinus entre les vecteurs qui les représentent, sachant que plus un mot est sémantiquement proche d'un autre, plus leurs vecteurs pointent dans la même direction et donc plus le cosinus se rapproche de 1. Ainsi, la valence d'un mot est calculée par la soustraction de la somme des cosinus entre ce mot et les points de repère positifs et celle des cosinus entre ce mot et les points de repère négatifs (Bestgen, 2006).

L'efficacité de la méthode a été évaluée grâce à une comparaison avec le General Inquirer Lexicon (Stone et al., 1966), contenant plus de 3000 mots étiquetés négatifs ou positifs. La valence de 64% des mots a été déterminée correctement (Bestgen, 2006).

- **Bestgen** (2002) a développé la procédure **DI-ASL**, qui, de la même façon que la PR-ASL, se base sur l'ASL pour déterminer la proximité entre les mots cibles et des mots germes dont la valence est connue. A la différence de la méthode de Turney et Littman (2002, 2003), le DI-ASL n'emploie pas 14 mots germes fixes comme références, mais un ensemble de points de repère sélectionnés pour chaque mot cible parmi un dictionnaire de 3000 mots dont la valence a été évaluée par des juges. Avec cette méthode, les 30 mots ayant le plus grand cosinus avec le mot cible sont identifiés à l'aide d'une ASL. La valence moyenne de ces 30 plus proches voisins, dont la valence est connue, est ensuite calculée. Cette dernière correspond à la valence du mot à analyser. Il est important de noter que la valence moyenne n'est pas pondérée par la proximité (Bestgen, 2006 ; Vincze & Bestgen, 2011).

Une comparaison entre les valeurs prédites et réelles pour 60 mots du dictionnaire a permis d'obtenir des corrélations entre 0,56 et 0,7.

A la différence des travaux cités ci-dessus, **Kamps et al. (2004)** ont mis au point deux méthodes qui visent, quant-à-elles, à estimer la valence sur la base de liens synonymiques.

- **KA1** est une méthode consistant à mesurer la distance géodésique (distance minimale) entre le mot cible et les mots germes « good » et « bad ». Ainsi, pour chaque mot à analyser, deux distances sont considérées : la distance la plus courte au « good » et la distance la plus courte au « bad ». Les mots se regroupent facilement en fonction de la longueur des chemins minimaux vers les adjectifs « good » et « bad » (Figure 2).

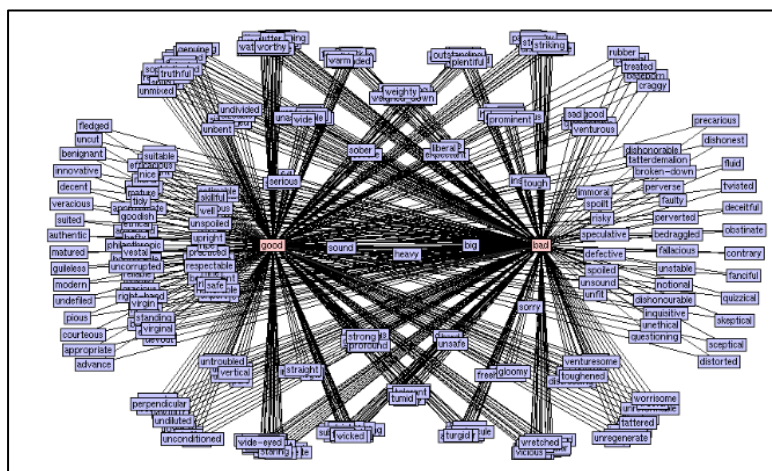


Figure 2 : Les distances géodésiques aux adjectifs "bon" et "mauvais" (Kamps et al., 2004)

La valence du terme cible est égale à sa distance relative par rapport aux deux références :

$$KA1(t) = \frac{d(t,mauvais) - d(t,bon)}{d(bon,mauvais)}$$

avec $d(i,j)$ la distance du plus court chemin synonymique entre les mots i et j

La différence des longueurs des chemins minimaux vers les deux mots de référence « good » et « bad » dépend de la distance entre ces deux mots. Par conséquent, il convient de diviser la différence $d(t, mauvais) - d(t, bon)$ par la distance entre « good » et « bad » afin d'obtenir une valeur dans l'intervalle $[-1,1]$. Ainsi, pour chaque mot, la fonction KA1 attribue une valeur allant de -1 (pour les mots du « mauvais » côté du lexique) à 1 (pour les mots du « bon » côté du lexique). Par exemple, pour l'adjectif « honnête », la fonction KA1 ressort la valeur 1 (Kamps et al., 2004).

$$KA1(honnête) = \frac{d(honnête,mauvais) - d(honnête,bon)}{d(bon,mauvais)} = \frac{6-2}{4} = 1$$

- **KA7** est une adaptation de KA1, qui prend en compte sept paires de référence : bon / mauvais, gentil / méchant, excellent / médiocre, positif / négatif, heureux / malheureux, correct / faux, supérieur / inférieur. Le nombre de paires de référence a ainsi été multiplié par sept par rapport à KA1. Il existe une fonction d'évaluation de la valence, reprenant la somme des évaluations pour chaque paire : (Vincze & Bestgen, 2011)

$$KA7(t) = \frac{\sum_{k=1}^n d(t, i_k) - \sum_{k=1}^n d(t, j_k)}{\prod_{k=1}^n d(i_k, j_k)}$$

où i_k et j_k forment une paire d'adjectifs des 7 paires prises en compte

Une étude a été réalisée grâce au calcul du pourcentage des termes bien classés pour les mots de General Inquirer (GI) (Stone et al., 1966) traités par toutes les méthodes (PR-ASL, DI-ASL, KA1 et KA7). Ainsi, le Tableau 3 souligne la supériorité de KA7 sur les autres méthodes. Cependant, la méthode KA7 concerne moins de mots que les méthodes basées sur l'ASL (688 mots du GI contre 1992). Pour un résultat très similaire, la méthode KA1 est moins chronophage et moins coûteuse que KA7 (Vincze & Bestgen, 2011).

<i>N</i>	<i>SO-ASL</i>	<i>DI-ASL</i>	<i>KA1</i>	<i>KA7</i>
550	64%	70%	80%	83%

Tableau 3 : Pourcentage de classification correcte pour les mots de GI traités par les méthodes (Vincze & Bestgen, 2011)

d) Logiciels de mesure de la valence d'un texte

Plusieurs logiciels développés pour l'évaluation émotionnelle de textes se basent sur l'hypothèse que la valence d'un texte peut être déterminée à partir de la valence de chaque mot qui le constitue. Ainsi, ces derniers utilisent des ressources lexicales donnant de l'information sur la polarité des mots telles que SentiWordNet (Esuli & Sebastiani, 2006). Ces logiciels diffèrent simplement par leur mode de détection des mots et leur classification. Si le texte contient plus de mots négatifs que positifs, alors la valence globale sera en conséquence négative (Ding et al., 2008).

- **SentiWordNet (SWN)** : SentiWordNet est une ressource lexicale assignant à chaque synset de la base de données WordNet trois valeurs, un synset étant un ensemble d'un ou plusieurs synonymes, un groupe de mots interchangeables dénotant un sens ou un usage particulier. Le fonctionnement de SentiWordNet nécessite l'inclusion de mots germes, connus comme étant

positifs, négatifs et objectifs. Lorsqu'un synset est analysé, SentiWordNet lui attribue un triple score de polarité (score de négativité, de positivité et d'objectivité) de façon à ce que la somme de ces trois scores soit égale à 1 (Figure 3) (Baccianella et al., 2009).

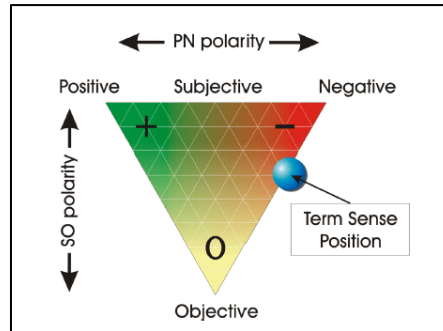


Figure 3 : Exemple de la représentation des termes après analyse par SentiWordNet

Le Tableau 4 résume les scores de l'adjectif « estimable » pour ses trois sens ; alors que le sens « calculable » est neutre, les deux autres sens sont très positifs (Chaumartin, 2012).

	P = 0, N = 0, O = 1	{computable#1, estimable#3} Maybe be computed or estimated ; « a calculable risk » ; « computable odds » ; « estimable assets »
	P = 0,75, N = 0, O = 0,25	{estimable#1} Deserving of respect or high regard
	P = 0,625, N = 0,25, O = 0,125	{honorable#5, good#4, respectable#2, estimable#2} Deserving of esteem and respect, « all respectable companies give guarantees » ; « ruined the family's good name »

Tableau 4 : Valence affective des trois sens de l'adjectif « estimable » selon SentiWordNet
(Chaumartin, 2012)

e) Mesure des émotions et produits

L'étude des émotions ne se limite pas aux corpus de textes, paragraphes, phrases et mots. Il existe également des méthodes développées pour mesurer les émotions suscitées par un produit. La plupart se concentre sur des mesures explicites d'auto-évaluation, demandant aux consommateurs d'indiquer leurs émotions lorsqu'ils consomment, sentent ou voient des produits (Cassini et al., 2020).

- **EsSense Profile (ESP)** est une méthode de mesure des émotions lors de l'évaluation d'un produit développée par King et Maiselman (2010). Cette méthode fournit une liste détaillée de 39 attributs émotionnels que les consommateurs associent aux produits testés. Cette liste peut être étendue ou modifiée pour tenir compte des émotions spécifiques qui peuvent être appropriées dans certaines catégories de produits. La collecte des données peut se faire soit en choisissant toutes les réponses applicables (CATA), soit en évaluant les 39 mots émotionnels sur une échelle d'intensité en 5 points allant de « pas du tout » à « extrêmement ». Les deux approches fournissent des informations utiles, cependant, l'approche par échelle fournit plus de détails, notamment lors de la comparaison de produits présentant de faibles différences (King et al., 2010). Malgré ses avantages, le EsSense Profile présente des limites ; lorsque Jaeger et al. (2013) ont demandé aux participants d'utiliser le EsSense Profile et de partager leurs pensées à voix haute, plusieurs participants ont indiqué qu'ils ne comprenaient pas certains mots d'émotion ou qu'ils les percevaient comme trop similaires. En outre, certains participants ont eu l'impression de ne pas ressentir fortement les mots d'émotion, que la tâche était longue et intimidante, et qu'ils se sentaient poussés par la procédure de test à varier leurs réponses et à indiquer leurs sentiments plus fortement (Kenney & Adhikari, 2016).

Une version abrégée de l'ESP, appelée EsSense 25 (compilant 25 termes émotionnels) a été développée par la suite (Nestrud et al., 2016).

EsSense Profile a été appliqué à différentes catégories d'aliments et de boissons comme le chocolat (Cardello et al., 2012), le jus de cassis (Ng et al., 2013) et la pâte à tartiner aux noisettes (Spinelli et al., 2015).

- **Geneva Emotion and Odor Scale (GEOS)** (Chrea et al., 2009) a été développée pour mesurer les sentiments affectifs qui renvoient à la stimulation olfactive dans une population francophone. GEOS contient 36 termes d'émotion classés en six dimensions: sensualité, relaxation, sentiment agréable, rafraîchissement, plaisir sensorielle et sentiment désagréable. Une étude a montré que l'utilisation des termes de GEOS permet d'obtenir de meilleures réponses pour les stimuli olfactifs que celle des termes des émotions de bases ; les participants étant plus en accord dans leurs réponses et les réponses étant plus discriminantes. Afin de simplifier l'outil GEOS, le lexique émotionnel a été réduit à 18 termes. Ce nouvel outil, nommé ScentMove, présente des résultats fortement corrélés à ceux obtenus lors de l'utilisation de GEOS (Porcherot et al., 2010).
- **Product Emotion Measurement Instrument (PrEmo)** est une méthode d'auto-évaluation graphique créée spécialement pour étudier les émotions suscitées par l'interaction avec les

produits. Le fonctionnement de cette méthode se base sur l'animation d'un personnage de cartoon, exprimant un ensemble de 14 émotions (7 positives, 7 négatives) (Figure 4) . Les participants évaluent leur ressenti pour chacune des émotions sur une échelle de 5 intensités (de « I do feel it strongly » à « I do not feel this ») (Desmet et al., 2000 ; Laurans & Desmet, 2012).

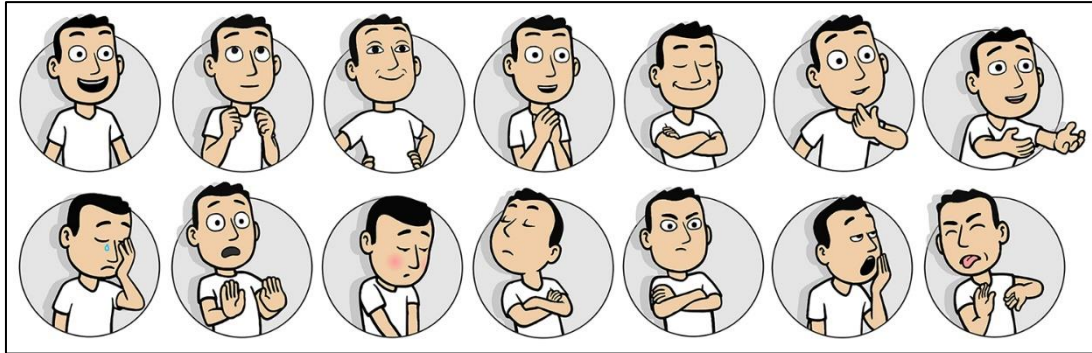


Figure 4 : Echelle PrEmo

Ces méthodes, développées dans le but de permettre aux participants de décrire plus facilement leurs émotions, présentent des avantages. D'après Dalenberg et al. (2014), les émotions, lorsqu'elles sont exprimées visuellement, sont supposées se rapprocher d'avantage des émotions vécues intuitivement, le traitement des émotions étant plus rapide pour les expressions faciales que pour les mots émotionnels. De plus, elles sont faciles à mettre en œuvre avec des enfants (Laurans & Desmet, 2012) et s'adaptent à différentes cultures en évitant la traduction des descripteurs d'émotion (Mojet et al., 2015).

- **Free Choice Profiling** (FCP) est une méthode où les évaluateurs décrivent les produits dans leurs propres mots (Kaneko et al., 2018). Elle vise ainsi à recueillir toutes les émotions qui viennent à l'esprit des participants concernant un produit donné suite à une étude visuelle, gustative, olfactive etc. La pertinence d'un terme vis-à-vis du produit étudié est corrélée à la fréquence à laquelle ce dernier est déclaré par les consommateurs. Cette méthode a notamment été utilisée pour identifier les caractéristiques des emballages de produits alimentaires qui attirent le plus l'attention des consommateurs (Ares & Deliza, 2010).

Une étude de Jaeger et al. (2013) a montré que près de 40% des termes cités à l'aide d'une déclaration libre ne sont pas des termes inclus dans le EsSense Profile. Ces résultats évoquent alors l'idée que les lexiques prédéfinis ne sont pas capables de générer toutes les émotions que les consommateurs éprouvent vis-à-vis d'un produit. Dans l'étude, les participants ont utilisé

en moyenne 27 termes émotionnels avec le EsSence Profile, contre 4 avec la déclaration libre. Cette différence trouve son origine dans deux principales hypothèses : les questionnaires prédéfinis aident les participants dans la verbalisation de ce qu'ils ressentent ; les termes émotionnels provenant du EsSence Profile sont utilisés en créant des associations avec le produit testé mais ne seraient pas de réelles émotions ressenties (Ng et al., 2013).

La liste de questionnaire présentée ci-dessus n'est pas exhaustive et pourrait être complétée par les questionnaires suivants :

- Profil EmoSensoriel : combinaison de 14-17 termes émotionnels et 13 termes sensoriels (Schouteten et al., 2015)
- Echelle ASR (Affect Self Report) : classement de 18 termes affectifs sur une échelle de 7 points (Christie & Friedman, 2004)
- EFT (Test alimentaire empathique) : classement de 12 termes empathiques (Geier et al., 2016)
- PANAS, tableau des affects positifs et négatifs : classement de 10 termes émotionnels positifs et 10 termes émotionnels négatifs (Watson et al., 1988).

II. Compréhension

a) Définition

La compréhension correspond à l'action de comprendre et à la faculté, la capacité de saisir et d'assimiler les choses. Elle peut également se définir comme la saisie intellectuelle du rapport de signification qui existe entre tel signe et la chose signifiée. La compréhension est constituée de la cognition, processus par lequel les entrées sensorielles sont transformées, et de l'affect, expérience des sentiments et émotions (CNRLT, s.d.).

La compréhension n'est pas un processus binaire, les degrés de compréhension sont plus ou moins profonds. Elle varie en fonction des individus, de leurs connaissances, expériences et aptitudes. La compréhension repose sur un ensemble de fonctions cognitives et exécutives : la langage, la mémoire, le raisonnement et les gnosies.

De nombreuses études ont été menées afin de saisir la compréhension des individus vis-à-vis de différents termes. Ces dernières peuvent être implicites, en faisant appel à l'inconscient, ou explicites, à l'aide d'instruments d'auto-évaluation.

b) Compréhension et terminologie sensorielle de texture

Au cours des dernières années, plusieurs méthodologies de collecte d'informations sur la compréhension et la définition qu'ont les consommateurs des descripteurs sensoriels ont été

développées. Ces dernières concernent particulièrement la texture, une propriété sensorielle complexe impliquant plusieurs paramètres très différents et contribuant à l'acceptabilité globale d'un aliment, ainsi qu'à son apparence et à sa saveur.

- **Varela et al. (2007)** ont étudié les différences dans les définitions et les termes utilisés par les consommateurs pour décrire leurs sensations associées au croustillant et au croquant. Pour cela, un questionnaire a été mis au point et présenté à 200 consommateurs de deux pays hispanophones différents. Ce dernier comportait sept questions ouvertes, dont trois sur les aliments croustillants (« crujiertes » en espagnol), trois sur les aliments croquants (« crocantes » en espagnol) et une sur les différences entre les deux termes (Figure 5).

INSTRUCTIONS:

Please answer the following questions in order.
Do not ask others, we are interested in YOUR opinion.
There are NO right or wrong answers

(1) Please mention 3 crispy food items:

(2) Define "crispy" in your own words:

(3) Imagine you are eating one of the food products you mentioned in (1).
Please describe the sensations you feel in your mouth. What happens when chewing?

(4) Now please mention 3 crunchy food items:

(5) Define "crunchy" in your own words:

(6) Imagine you are eating one of the food products you mentioned in (4).
Please describe the sensations you feel in your mouth. What happens when chewing?

(7) In your opinion, are there differences between "crispy" and "crunchy" foods? If any, please describe them.

Figure 5 : Questionnaire de compréhensions issu des travaux de Varela et al. (2008)

Des analyses du Chi carré des comparaisons entre pays et entre attributs ont été effectuées pour déterminer si les fréquences de réponse observées différaient de manière significative. Les réponses ont montré que les termes croustillant et croquant avaient des significations

différentes ou évoquaient des perceptions différentes selon le pays. La principale différence entre les deux pays est que 38% des consommateurs espagnols ne connaissent pas le terme « crocantes ».

- **Fillion et Kilcast** (2002) ont également contribué à une meilleure connaissance de la terminologie utilisée par les consommateurs pour décrire les sensations liées au croustillant et au croquant. Leur étude s'est concentrée sur les termes des consommateurs relatifs aux caractéristiques des fruits et légumes et a été réalisée à l'aide de la méthode Free Choice Profiling (FCP) auprès de consommateurs naïfs et de panélistes formés. Six produits ont été sélectionnés pour l'expérience (carotte, céleri, concombre, poivron vert, Granny Smith et Golden Delicious) de manière à présenter un large éventail de caractéristiques de texture.

Au total, 10 consommateurs et 11 panélistes formés ont participé à l'étude. Ils ont chacun passé un entretien individuel où des paires de fruits et légumes différents leur ont été présentées, de sorte que tous les produits soient vus deux fois dans des combinaisons différentes. Lors de cet entretien, ils ont été invités à décrire les similitudes et les différences de texture perçues, particulièrement au niveau de la sensation en bouche, de la décomposition de l'échantillon, des sons et de la mastication. Par la suite, un vocabulaire personnel, spécifique à chaque répondant, a été défini. Cette étape a été suivie d'une séance de notation, au cours de laquelle les répondants ont été invités à évaluer chaque produit sur une échelle d'intensité pour chaque attribut de leur propre glossaire.

Les deux ensembles de données ont été analysés à l'aide de l'Analyse Généralisée Pro-crustes (GPA), qui a permis de générer des cartes de produits et d'attributs. Les produits qui étaient proches les uns des autres sur la carte ont été interprétés comme étant similaires en termes de texture. De la même manière, les attributs proches les uns des autres sur la carte ont été interprétés comme étant similaires par les consommateurs. Les cartes obtenues via GPA ont montré que les perceptions sensorielles exprimées par un panel formé présentaient une bonne corrélation avec la terminologie des consommateurs.

Sur les dix consommateurs interrogés, neuf ont utilisé le terme « crunchy » pour décrire la texture des produits, mais cinq seulement ont mentionné le terme « crispy ». Ceux qui ont mentionné les deux ont toujours utilisé les deux termes avec des significations différentes. Le terme « crispy » désigne une texture légère et fine produisant une rupture nette et un son aigu lorsqu'une force est appliquée, principalement lors de la première bouchée avec les incisives. En revanche, le terme « crunchy » serait associé à une texture dure et dense qui se fracture sans

déformation préalable, produisant un son fort et grave qui est répété à plusieurs reprises lors de la mastication.

Les méthodes présentées ci-dessus demandent aux participants de mettre des mots sur ce qu'ils comprennent des descripteurs sensoriels de texture ou sur les sensations qu'ils ressentent. Or, il est difficile pour les consommateurs de transformer leur perception sensorielle en une description intelligible. D'autres approches, telles que celle de Bhushan et al. (1997), ont permis d'accéder indirectement à la compréhension des individus.

- Bhushan et al. (1997) ont développé une approche faisant appel à un système de comparaisons afin de comprendre la catégorisation visuelle des termes sensoriels. Ils ont accédé aux représentations mentales visuelles des termes sensoriels des individus grâce à une tâche de tri libre. 98 mots de texture ont été sélectionnés en fonction de leur fréquence d'apparition et imprimés sur des cartes. Les répondants ont été invités à visualiser la surface correspondant à chaque mot et à regrouper les mots si leurs surfaces visualisées apparaissaient similaires. Les sujets ne disposaient d'aucun critère explicite pour juger de la similarité et pouvaient créer autant de groupes qu'ils le souhaitaient. Une fois les groupes terminés, chaque participant a décrit chaque groupe et expliqué pourquoi les éléments avaient été jugés similaires. Le traitement des données a été réalisé à l'aide de la MDS et d'une classification hiérarchique.

Dans un second temps, Bhushan et al. (1997) ont analysé la correspondance entre images et mots de texture afin de répondre aux questions suivantes : « quels types d'images de texture ces mots évoquent-ils ? », « quels types de mots sont associés à cette image de texture ? ». Pour cela, 20 participants ont été divisés en deux groupes. L'un avait pour consigne de trouver le ou les mot(s) correspondant à une image et l'autre de trouver la ou les image(s) correspondant à un mot. 56 images de textures étaient à disposition des participants. Le traitement des données a été réalisé à l'aide d'un tableau de contingence puis du calcul d'un coefficient de contingence de Pearson pour les cartes image-mot et mot-image.

c) Compréhension et allégations de santé

En réponse aux préoccupations croissantes des consommateurs vis-à-vis de leur santé, les fabricants de produits communiquent de plus en plus à l'aide d'allégations de santé, c'est-à-dire toute allégation qui « affirme, suggère ou implique l'existence d'une relation entre une catégorie de denrées alimentaires, une denrée alimentaire ou l'un de ses composants et la santé » (règlement (CE) n° 1924/2006). Ces dernières sont censées faciliter des choix éclairés et promouvoir une alimentation

saine. Cependant, en réalité, la compréhension de ces allégations par les consommateurs n'est pas toujours simple. Plusieurs méthodes de mesure de la compréhension des allégations de santé ont été développées, allant de méthodes qualitatives (entretiens approfondis, analyse linguistique...), à des méthodes mixtes et à des méthodes quantitatives. En règle générale, les consommateurs semblent préférer les allégations de santé plus courtes, exprimées en langage normal, plutôt que les allégations longues et complexes en langage scientifique.

- **Grunert et al** (2011) ont mesuré la compréhension d'une allégation de santé par 720 consommateurs à l'aide de la méthodologie **Consumer Understanding Test** (CUT) développée par Danone (Rogeaux, 2010). Cette dernière consiste en la présentation d'une allégation de santé aux participants (ici « Actimel aide à renforcer les défenses naturelles de l'organisme »), accompagnée de l'emballage du produit et/ ou d'une publicité télévisée dans lequel le produit apparaît. Deux questions ouvertes leur ont ensuite été posées :
 1. Après avoir vu ce pack et/ ou cette publicité, si vous deviez dire à un ami ce que fait XXX, que diriez-vous?
 2. Et si vous deviez dire à un ami comment ça marche?

Le contenu des réponses a été analysé dans un schéma de codage hiérarchique, et chaque code résultant catégorisé comme sûr (déclaration conforme), risqué (déclaration non conforme) ou vague (déclaration vague ou non pertinente au regard de l'allégation de santé) en comparaison avec les bases scientifiques de l'allégation.

- **Stancu** (2016) a mesuré la compréhension de trois allégations de santé relatives à l'huile d'olive :
 1. « 2 cuillères à soupe (environ 20 g) d'huile d'olive par jour peuvent réduire le risque de maladie coronarienne » (allégation non autorisée dans l'UE mais approuvée aux États-Unis)
 2. « Les polyphénols de l'huile d'olive contribuent à la protection des lipides sanguins contre le stress oxydatif. L'effet bénéfique est obtenu par une consommation quotidienne de 20 g d'huile d'olive » (allégation autorisée dans l'UE)
 3. « 20 g de cette huile d'olive par jour aident à protéger les lipides sanguins contre les réactions nocives » version de l'allégation de santé 2, mise au point aux fins de l'étude

L'échantillon était composé de 2024 personnes habituées à utiliser de l'huile d'olive et âgés de 18 à 74 ans. Parmi les 2024 répondants, 1000 étaient italiens et 1024 danois (ces deux pays ayant été sélectionnés en raison de la différence de la place de l'huile d'olive dans leur régime alimentaire).

Avant de demander aux participants d'évaluer plusieurs affirmations à l'aide d'une échelle, Stancu (2016) a utilisé la méthode CUT sur les trois allégations afin de pouvoir comparer les résultats entre les deux méthodes. Les participants ont ainsi reçu l'instruction d'imaginer une huile d'olive extra vierge portant l'une des trois allégations (sélectionnée selon un plan randomisé) et de répondre aux deux questions ouvertes de la méthode.

La deuxième mesure de la compréhension a ensuite consisté à évaluer des déclarations préformulées reflétant les significations potentielles des allégations de santé. Elle a toujours été réalisée après la mesure CUT afin d'éviter de biaiser les réponses aux questions ouvertes. Les participants ont vu la même allégation de santé dans les deux mesures. Chaque répondant devait indiquer dans quelle mesure il était d'accord ou non avec les 15 affirmations, affichées dans un ordre aléatoire, sur une échelle de 7 points d'accord allant de "pas du tout d'accord" à "tout à fait d'accord". Les affirmations ont été élaborées au préalable en collaboration avec des nutritionnistes afin que ces dernières reflètent objectivement des inférences sûres (conformes aux preuves scientifiques de l'allégation) ou risquées (non conformes aux preuves scientifiques de l'allégation). Les mêmes affirmations ont été utilisées pour les trois allégations. Les réponses obtenues ont permis de calculer un indice de compréhension à l'aide de la formule suivante : $IC = \text{score moyen pour les items sûrs} - \text{score moyen pour les items risqués}$. Cet indice reflète ainsi la différence entre l'accord/ le désaccord des consommateurs avec les inférences sûres en moyenne et leur accord/leur désaccord avec en moyenne, ce qui donne une indication de la capacité des gens à faire la différence entre les items sûrs et les items risqués. Plus l'indice de compréhension est élevé, plus la compréhension est bonne.

Dans l'ensemble, la méthode CUT a permis de mieux saisir les différences de compréhension des consommateurs entre les allégations de santé. Il existe un lien entre la mesure CUT et la mesure de l'indice de compréhension, ce qui montre une validité convergente. Les consommateurs classés comme sûrs ou risqués ont obtenu un score plus élevé sur l'indice de compréhension par rapport aux consommateurs classés comme autres.

d) Compréhension et descripteurs du vin

De nombreux descripteurs spécifiques sont utilisés pour communiquer sur les propriétés sensorielles des vins. Alors que certains sont unidimensionnels et univoques, d'autres sont multidimensionnels et équivoques tels que minéral, doux ou encore complexe. Ces dernières années, plusieurs études ont été menées afin de comprendre la signification de ces concepts mal définis.

- **Le Fur et al.** (2011) ont étudié la compréhension des concepts de vin gras et de vin ayant du volume par un panel d'experts et un panel de non professionnels. Pour chaque concept étudié, des tests d'association libre, de définition et de synonymie sous forme de questionnaires ont été réalisés. Le test d'association visait à étudier les mots spontanément associés aux concepts étudiés à l'aide de la question suivante : « si l'on vous dit vin gras (ou vin qui a du volume), à quoi pensez-vous ? ». Pour le test de définition et de synonymie, les auteurs ont invité les participants à donner une définition sous la forme de phrases courtes et à lister un ensemble de synonymes dans le contexte suivant : « Imaginez que vous ayez à expliquer à un ami ce qu'est un vin gras (ou un vin qui a du volume). Pour lui expliquer, vous pouvez en donner une définition et citer des synonymes ».

Les participants ont ainsi généré librement du vocabulaire, qui a été simplifié et harmonisé (regroupement des mots de même origine sémantique - verbe / adverbe, nom / adjectif, masculin / féminin - et de certaines synonymies franches). Les jeux de données ainsi obtenus ont été analysés grâce à une analyse de fréquence d'occurrence pour chacun des termes. Cela a permis de connaître le pourcentage de répondants ayant cité le terme et de déterminer quels étaient les termes les plus fréquemment employés pour qualifier les concepts de gras et de volume. L'ordre de citation des mots a également été pris en compte.

Les résultats obtenus ont permis aux auteurs de proposer une définition de chacun des deux concepts. Il convient de remarquer que les réponses des professionnels sont plus riches et plus pertinentes que les réponses de non-professionnels car moins dispersées.

- **Rodrigues et al.** (2015) ont développé une méthode afin d'étudier la compréhension du concept de minéralité. Le panel était constitué de 47 consommateurs de vin et de 40 vignerons de Chablis. Dans le cadre de l'étude, les participants ont été invités à décrire tous les mots qui leurs venaient à l'esprit lorsqu'ils entendaient le mot « minéralité » puis à évaluer l'importance de chaque mot produit sur une échelle en dix points (allant de 1 (pas important) à 10 (très important)). La valence de chaque mot était ensuite à évaluer sur une échelle en sept points (-

3 → + 3). Les données ont été soumises à une analyse prototypique afin d'identifier les domaines centraux et périphériques de la représentation sociale des experts et des consommateurs. Les résultats ont montré que les viticulteurs et les consommateurs de vin ont une manière différente de conceptualiser la minéralité.

Les travaux de Rodrigues et al. (2015) ont été inspirés de ceux de **Parr et al.** (2011), qui ont étudié le mot « complexe » grâce à une tâche de verbalisation réalisée par 69 participants (39 professionnels et 30 consommateurs). Lors des entretiens, les participants ont été invités à indiquer tous les mots que leur évoquait le concept de complexité puis à définir ce qu'ils entendaient par chaque mot produit. Ils ont ensuite dû classer ces mots par ordre d'importance / de priorité et évaluer leur caractère positif ou négatif sur une échelle en sept points. Les données obtenues ont été traitées à l'aide d'un logiciel permettant l'analyse de données textuelles afin de déterminer les termes les plus utilisés et donc de définir la complexité.

e) Compréhension et terminologie médicale

La communication entre le médecin et le patient au cours d'une consultation est un processus complexe qui comporte plusieurs éléments, notamment le vocabulaire utilisé et le ton de la voix. La compréhension de la terminologie médicale par les patients est fondamentale pour comprendre la nature de la maladie/des symptômes et le processus de diagnostic, ainsi que pour prendre des décisions éclairées concernant les soins et le suivi du traitement (O'Connell et al., 2013). Ainsi, de nombreuses études visant à déterminer si les participants ont compris le concept étudié et sont capables d'en donner une définition claire ont été menées.

- **Gibbs** (1986) a développé une méthode de mesure de la compréhension des termes médicaux, impliquant la participation de 50 patients du Primary Care Center de l'hôpital NewHaven. Les 15 termes dont la compréhension a été évaluée ont été sélectionnés dans des brochures d'information disponibles au Primary Care Center ; il s'agit de mots d'application universelle (abdomen, bouche, infection, symptôme...) ou se référant à une maladie prévalente (allergie, diabète, accident vasculaire cérébral...). Les participants à l'étude ont été invités à définir les 15 termes médicaux au cours d'un entretien d'une quinzaine de minutes avec le chercheur. Pour chaque terme, l'auteur énonçait le mot à définir et clarifiait le contexte grâce à une phrase typique (identique pour chaque patient testé). Le patient était ensuite encouragé à définir le terme verbalement. Les réponses ont été enregistrées au fur et à mesure par écrit sur un formulaire complété par l'expérimentateur.

Par la suite, les définitions ont été évaluées par un médecin n'ayant pas participé aux entretiens et classées en trois catégories :

- A. Les réponses qui révèlent une bonne compréhension du mot sont classées dans la catégorie A. Il n'est pas attendu qu'elles correspondent parfaitement à la définition restrictive du dictionnaire, mais qu'elles montrent que le patient comprend le mot dans le contexte de la phrase d'exemple.
- B. Les réponses qui révèlent une compréhension vague et partielle sont classées dans la catégorie B.
- C. Les réponses qui révèlent une compréhension incorrecte ou une absence de connaissance du mot sont classées dans la catégorie C.

D'autres études visant à mesurer la compréhension de termes médicaux grâce à une méthodologie similaire à celle de Gibbs ont été menées. Parmi ces dernières, celle de **O'Connell et al.** (2013), étude visant à évaluer la connaissance des définitions des termes communs dans la clinique du sein, peut être citée. Il a été demandé à 102 participants de compléter les définitions de 12 termes grâce à un questionnaire dans la salle d'attente avant leur consultation. Un score a ensuite été attribué à chaque réponse par deux des expérimentateurs en fonction de leur exactitude (0 : aucune réponse/ réponse incorrecte ; 1 : partiellement correcte ; 2 : complètement correcte).

- **Hadlow et Pitts** (1991) ont étudié la compréhension des termes médicaux à l'aide d'un questionnaire à choix multiples de 12 questions, dont 2 de médecine générale, 3 de neurologie et 7 de psychologie. Pour chaque terme, quatre définitions simplifiées par les chercheurs ont été proposées. Les participants ne pouvaient sélectionner qu'une unique définition par question et indiquer s'ils étaient sûrs ou non de leur réponse. L'échantillon était composé de 40 médecins, 40 infirmières, 20 travailleurs de soutien sanitaire et 120 patients ; le panel étant ainsi formé par des individus dont le niveau de connaissance était très variable dans le domaine.

Les niveaux de compréhension ont été définis en fonction du pourcentage de réponses correctes : « niveau élevé de compréhension » si 70% ou plus des réponses étaient correctes, « niveau moyen de compréhension » si 40-69% des réponses étaient correctes, « niveau bas de compréhension » si 39% ou moins des réponses étaient correctes.

- **Lerner et al** (2000) ont développé une méthode afin d'étudier la compréhension des termes médicaux par les patients des services d'urgence. Le questionnaire proposé à 249 répondants

comportait 21 questions basées sur quatre paires de mots analogues : cassé / fracturé, saignement / hémorragie, crise cardiaque / infarctus du myocarde, et sutures / points de suture ; et deux paires de mots non analogues : diarrhée / selles molles et plâtre / attelle. Les participants ont répondu à trois questions pour chaque paire de mots : "Y a-t-il une différence entre les termes ?", "Quel est le pire ?" et "Définissez la différence entre les termes". S'ils répondaient qu'il y avait une différence entre les termes, les autres questions relatives à ces termes n'étaient pas posées.

Un score a été déterminé pour chaque participant en fonction de l'identification correcte des paires de mots analogues. Des analyses univariées ont été effectuées en calculant les coefficients de corrélation de Pearson pour les variables continues et un t-test pour les variables dichotomiques.

- **Calamusa et al.** (2012) ont construit un questionnaire visant à évaluer la compréhension et les connaissances des consommateurs concernant l'automédication. Leur enquête, menée sur 1206 adultes, a été divisée en trois sous-sections. La première (vocabulaire) a étudié la compréhension des 17 termes les plus fréquents dans les notices d'informations. Pour cela, il a été demandé aux participants de placer 12 termes dans la section correcte d'un corps stylisé divisé en quatre sections et de répondre à un questionnaire à choix multiples pour 5 termes (« analgésique », « principe actif », « dosage », « contre-indications » et « interactions »).

La deuxième sous-section a évalué la capacité à effectuer des calculs de dosages simples, en demandant aux participants de calculer le nombre maximum de comprimés à ne pas dépasser par jour sur la base des instructions "prendre 1 ou 2 comprimés une ou deux fois par jour". Enfin, la dernière sous-section a quantifié les connaissances sur la signification, la distribution et les symboles d'identification des médicaments.

f) Autres études de la compréhension

De nombreuses études ont été menées afin d'évaluer la compréhension de divers sujets tels que les concepts de base de la physique-chimie ou encore de l'assurance maladie.

- Une enquête a été menée par Loewenstein et al (2013) afin de mesurer la compréhension des concepts de base de l'assurance maladie. 202 participants ont répondu à l'étude ; ces derniers étant des adultes non institutionnalisés âgés de 25 à 64 ans résidant aux États-Unis, les principaux décideurs pour leurs propres soins de santé ou ceux de leur famille ; et ayant une

assurance maladie par l'intermédiaire de leur propre employeur ou de celui d'un membre de leur famille.

La première partie de l'étude a visé à déterminer si les répondants comprenaient chacun des 4 concepts fondamentaux de l'assurance : franchise, copaiement, coassurance et contribution maximale. Elle s'est déroulée en deux étapes : une question fermée afin de leur demander s'ils savaient ce que le terme signifie ; puis une question à choix multiples afin de vérifier s'ils comprenaient réellement le terme (Figure 6).

1. *Savez-vous ce qu'est une franchise ?*

Oui

Non

2. *Laquelle des propositions suivantes décrit le mieux une franchise ?*

Un montant déduit de votre salaire pour payer votre prime d'assurance

Le montant déduit du total de vos frais médicaux annuel

Le montant que vous payez avant que votre compagnie d'assurance ne verse des prestations

Le montant que vous payez avant que vos frais de santé ne soient entièrement couverts

Je ne suis pas sûr

Figure 6: Questionnaire de compréhensions issu des travaux de Loewenstein et al (2013)

Une police d'assurance conventionnelle, comprenant des franchises, des copaiements, des coassurances et des contributions maximales, a ensuite été présentée aux personnes interrogées. Ces dernières ont d'abord été invitées à imaginer qu'elles étaient assurées au titre de la police qui leur était présentée, puis à répondre à une série de questions à choix multiples sur les coûts des services médicaux dans différents scénarios. Les réponses à ces questions ont été sélectionnées de manière à ce qu'elles ne nécessitent que peu de calculs mais qu'elles exigent, et donc mesurent, la compréhension des mécanismes de l'assurance maladie. Après chacune de ces questions, une question ouverte leur a été posée pour savoir s'il était difficile de répondre à la question à choix multiples et, si oui, pourquoi.

Enfin, les dernières questions ont permis de quantifier leur compréhension de différents concepts et questions à l'aide d'une échelle en 5 points allant de « ne comprend pas du tout »

à « comprend vraiment » tels que « comment fonctionnent les franchises individuelles et familiales ? ».

g) Comparaison entre experts sensoriels et consommateurs

L'opposition entre les panels d'experts et de consommateurs n'est pas si franche. **Worch et al.** (2010) ont montré que les résultats entre les deux panels sont proches et que les cartes de l'espace produit générées sont similaires. **Ares et Varela** (2017), qui vont dans le même sens des travaux de Worch et al. (2010), précisent tout de même que le vocabulaire des consommateurs n'est pas aussi développé que celui des experts. Cette différence se ressent particulièrement dans les études sur le vin, où les experts sont détenteurs d'un langage spécialisé, difficile d'accès aux novices (« vin doux », « joli bouquet »...). Les descriptions réalisées par des experts comportent davantage de termes techniques et précis alors que celles des novices comportent davantage de termes globaux.

Langlois et al. (2001), qui ont comparé le lexique associé au « vin de garde » de quatre groupes de dégustateurs (professionnels, connaisseurs, consommateurs, panel entraîné), ont montré que le vocabulaire utilisé diffère en fonction du niveau d'expertise et de l'expérience. Les panélistes ont principalement utilisé une terminologie sensorielle normée, apprise lors de leurs séances d'entraînement.

III. Adéquation

a) Définition

Le terme adéquation désigne « la qualité de ce qui est exactement adapté, approprié au but visé » (CNRTL, s.d.). Il est employé pour décrire un rapport de convenance parfaite.

b) Adéquation des descripteurs sensoriels

- **Schutz** (1994) a détaillé une approche permettant de mesurer l'adéquation grâce à une échelle en sept points allant de « jamais approprié » à « toujours approprié ». En 2006, cette approche a permis à **Caporale et al.** (2006) d'évaluer la pertinence de certains attributs sensoriels dans la description des huiles d'olive de Ogliarola, Maiatica et Coratina. Les neuf descripteurs sensoriels de l'étude ont été sélectionnés parmi ceux utilisés le plus couramment dans l'évaluation sensorielle des huiles d'olive. Les 34 consommateurs ont été invités à évaluer l'adéquation des attributs pour décrire les propriétés sensorielles des huiles étudiées à l'aide de l'échelle en sept points, développée par **Schutz** (1994). Les résultats ont permis de sélectionner

les attributs « piquant » et « amertume », pour lesquels les notes moyennes d'adéquation étaient les plus élevées.

Cette échelle a également été utilisée pour mesurer l'adéquation des aliments à une variété d'usages et de situations tels que le moment de la journée ou encore le lieu où la nourriture est servie. C'est le cas de **Cardelo et Schutz** (1996), qui ont invités 300 participants à évaluer un échantillon (fruits, légumes, bonbons, jambon fumé...) en fonction de son adéquation avec dix situations différentes (avoir faim, manger seul, vouloir quelque chose de nutritif, avoir besoin de se remonter le moral...).

- **Cadena et al.** (2017) ont utilisé un questionnaire Check-All-That-Apply (CATA), où ils ont invité 243 consommateurs à sélectionner les termes qu'ils jugeaient appropriés pour décrire huit yaourts (Figure 7). Les 19 attributs, relatifs aux caractéristiques sensorielles des yaourts, ont été sélectionnés sur la base des données publiées, en tenant compte des descripteurs générés par les évaluateurs formés lors de l'analyse descriptive.

Cochez tous les attributs appropriés pour décrire l'échantillon :		
<input type="checkbox"/> Sucré	<input type="checkbox"/> Lisse	<input type="checkbox"/> Fluide
<input type="checkbox"/> Collant	<input type="checkbox"/> Sans saveur	<input type="checkbox"/> Grumeleux
<input type="checkbox"/> Hétérogène	<input type="checkbox"/> Epais	<input type="checkbox"/> Aigre
<input type="checkbox"/> Crémeux	<input type="checkbox"/> Saveur de lait	<input type="checkbox"/> Rugueux
<input type="checkbox"/> Liquide	<input type="checkbox"/> Doux	<input type="checkbox"/> Homogène
<input type="checkbox"/> Visqueux	<input type="checkbox"/> Corsé	<input type="checkbox"/> Consistant
<input type="checkbox"/> Saveur de vanille		

Figure 7 : Exemple du questionnaire CATA de l'étude de Cadena et al. (2014)

Toutefois, malgré le fait que la fréquence de mention des termes dans les questions CATA soit étroitement liée à l'intensité de l'attribut, elles ne fournissent pas d'informations quantitatives puisque les consommateurs évaluent uniquement si un terme est approprié ou non pour décrire le produit. Ainsi, la méthode RATA (Rate-All-That-Apply), variante du CATA, peut être

envisagé à la place. Lorsque l'on utilise le RATA, les mots appropriés sélectionnés sont notés sur une échelle d'intensité.

L'utilisation d'un RATA a été mise en œuvre par **Waehrens et al.** (2016) sur les descripteurs sensoriels du chocolat. Après élaboration de la liste des descripteurs par un panel externe formé, 11 participants ont été invités à cocher les attributs qu'ils jugeaient appropriés pour décrire le chocolat, et ensuite à évaluer leur intensité sur un échelle en trois points (« peu », « moyen », « beaucoup »).

Il pourrait également être intéressant de réaliser une CATA inversée. Après avoir pris connaissance des produits proposés, le consommateur serait invité à sélectionner tous les produits qui, selon lui, peuvent être caractérisés par un descripteur.

- **Rose et al.** (2004) ont développé une procédure adaptée aux enfants âgés de 6 à 11 ans afin de générer des descripteurs et de mesurer leur adéquation. 208 enfants ont été invités à décrire l'apparence, l'odeur, le goût et la sensation en bouche d'un produit à base de viande (saucisse, hamburger, côte d'agneau ou rumsteck). Les descripteurs générés par plus d'un enfant ont été inclus dans la phase suivante de l'étude : la mesure de l'adéquation des descripteurs à un produit. Lors de cette dernière, les auteurs ont demandé aux enfants, un par un, de répondre « oui » ou « non » à la question : « ce mot est-il approprié pour décrire le produit ? ». Afin de les familiariser avec cette tâche et de leur démontrer l'inadéquation de certains descripteurs, l'exemple de la pomme a été utilisée : « utiliser le mot « bleu » pour décrire la pomme n'est pas approprié ni adapté ».

c) Alternative : tâche de catégorisation

Afin de déterminer si un terme est adapté pour décrire du point de vue sensoriel la catégorie de produit, les consommateurs pourraient réaliser une tâche de catégorisation. Cette dernière permettrait de déterminer le degré d'appartenance du produit au concept sensoriel.

Dacremont (2013) suggère de demander aux participants d'indiquer si le produit correspond, ou non, à la représentation qu'ils se font du concept sensoriel étudié (dans son cas, la croustillance). Pour ce faire, quatre réponses possibles étaient proposées, à l'image d'une étude sur le cheddar (Dacremont & Vickers, 2007) :

- C'est croustillant, et j'en suis sûr
- C'est croustillant, mais je n'en suis pas sûr

- Ce n'est pas croustillant, et j'en suis sûr
- Ce n'est pas croustillant, mais je n'en suis pas sûr

Cette méthode semble adaptée à des sujets novices puisqu'il s'agit d'une tâche naturelle, qu'ils effectuent spontanément au quotidien. Une tâche de catégorisation est alors, à l'évidence, plus familière qu'une notation sur une échelle.

Bibliographie

- Al-Shabi, M. (2020). Evaluating the performance of the most important Lexicons used to Sentiment analysis and opinions mining. *IJCSNS International Journal of Computer Science and Network Security*, 20, 51-57.
- Ares, G., Deliza, R. (2010). Studying the Influence of Package Shape and Colour on Consumer Expectations of Milk Desserts Using Word Association and Conjoint Analysis. *Food Quality and Preference*, 21, 930-937.
- Ares, G., Varela, P. (2017). Trained vs. consumer panels for analytical testing: Fueling a long lasting debate in the field. *Food Quality and Preference*, 61, 79-86.
- Baccianella, S., Esuli, A., Sebastiani, F. (2010). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. *Lrec*. 2200-2204.
- Bestgen, Y. (2002). Détermination de la valence affective de termes dans de grands corpus de textes. *Actes du Colloque International sur la Fouille de Textes*, 81-94. Nancy: INRIA.
- Bestgen, Y. (2006). Déterminer automatiquement la valence affective de phrases : Amélioration de l'approche lexicale. *Actes des JADT*, 179-188.
- Blanc, N., Brouillet, D., Syssau, A. (2006). Emotion et cognition : quand l'émotion parle à la cognition. Paris: In Press. 218 p.
- Bradley, M.M., Lang, P.J. (1999). Affective norms for English words (ANEW) : Stimuli, instruction manual and affective ratings. The Center for Research in Psychophysiology, University of Florida.
- Bradley, M.M., Lang, P.J. (2000). Measuring emotion: Behavior, feeling and physiology. *Cognitive neuroscience of emotion*, 242-276. New York, USA.
- Bhushan, N., Rao, A.R., Lohse, G.L. (1997). The Texture Lexicon: Understanding the Categorization of Visual Texture Terms and Their Relationship to Texture Images. *Cognitive Science*, 21, 219-46.
- Cadena, R., Caimi, D., Jaunarena, I., Lorenzo, I., Vidal, L., Ares, G., Deliza, R., Giménez, A. (2014). Comparison of rapid sensory characterization methodologies for the development of functional yogurts. *Food Research International*, 64, 446 – 455.

- Calamusa, A., DiMarzio, A., Cristofani, R. (2012). Factors that influence Italian consumers' understanding of over-the-counter medicines and risk perception. *Patient Educ Couns*, 87, 395-401.
- Caporale, G., Policastro, C., Carlucci, A., Monteleone, E. (2006). Consumer expectations for sensory properties in virgin olive oils. *Food Quality and Preference*, 17, 116-125.
- Cardello, A., Meiselman, H., Schutz, H., Craig, C., Given, Z., Leshner, L., Eicher, S. (2012). Measuring emotional responses to foods and food names using questionnaires. *Food Quality and Preference*, 24, 243-250.
- Cassani, L., Fiszman, S., Alvarez, M., Moreira, M., Laguna, L., Tarrega, A. (2020). Emotional Response Evoked When Looking At And Trying A New Food Product, Measured Through Images And Words. A Case-Study With Novel Fruit And Vegetable Smoothies. *Food Quality and Preference*, 84, 103955.
- Chaumartin, F.R. (2012). Antelope, une plate-forme de TAL permettant d'extraire les sens du texte : théorie et applications de l'interface syntaxe-sémantique. *Informatique et langage*. Université Paris-Diderot.
- Chrea, C., Grandjean, D., Delplanque, S., Cayeux, I., Le Calvé, B., Aymard, L. (2009). Mapping the semantic space for the subjective experience of emotional responses to odours. *Chemical Senses*, 34, 49–62.
- Christie, I. C., Friedman, B. (2004). Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach. *Int. J. Psychophysiol.* 51, 143–153.
- CNRTL. (2012). Valence - Compréhension – Adéquation. In Centre national de ressources textuelles et lexicales. Disponible à l'adresse : <https://www.cnrtl.fr/>
- Dacremont, C. (2003). Croustillant : aspects méthodologiques de la mesure sensorielle d'une caractéristique de texture complexe. [en ligne]. [Consulté le 12/08/2020]. Disponible à l'adresse suivante : <http://journals.openedition.org/aof/1227>
- Dacremont, C., Vickers, Z.M. (2007). Classification of cheese according to their closeness to the Cheddar cheese concept. *Journal of Sensory Studies*, 9, 237-246.
- Dalenberg, J., Gutjar, S., Ter Horst G., De Graaf, K., Renken, R., Jager, G. (2014). Evoked emotions predict food choice. *PLoS one* 9.

- Desmet, P., Hekkert, P., Jacobs, J. (2000). When a car makes you smile: Development and application of an instrument to measure product emotions. *Advances in Consumer Research*, 27, 111–117.
- Ding, X., Liu, B., Yu, P.S. (2008). A holistic lexicon-based approach to opinion mining. 2008 international conference on web search and data mining, 231-240. Palo Alto, USA.
- Esuli, A., Sebastiani, F. (2006). SENTIWORDNET: A publicly available lexical resource for opinion mining. *LREC'06*, 417–422.
- Fillion, L., Kilcast, D. (2002). Consumer perception of crispness and crunchiness in fruits and vegetables. *Food Quality and Preference*, 13, 23–29.
- Geier, U., Büssing, A., Kruse, P., Greiner, R., Buchecker, K. (2016). Development and application of a test for food-induced emotions. *PLoS one* 11.
- Gibbs, R. (1989). Patient understanding of commonly used medical vocabulary. *Yale Medicine Thesis Digital Library*. 2644.
- Grunert, K.G., Scholderer, J., Rogeaux, M. (2011) Determinants of consumer understanding of health claims. *Appetite* 56, 269–277.
- Hadlow, J., Pitts, M. (1991). The understanding of common health terms by doctors, nurses and patients. *Soc Sci Med*, 32, 193-196.
- Jaeger, S., Cardello, A., Schutz, H. (2013). Emotion questionnaires: A consumer-centric perspective. *Food Quality and Preference*, 30, 229–241.
- Kamps, J., Marx, M., Mokken, R.J. (2004). Using WordNet to Measure Semantic Orientation of Adjectives. *LREC*, 5, 1115-1118
- Kaneko, D., Toet, A., Brouwer, A.M., Kallen, V., Van Erp, J. (2018). Methods for Evaluating Emotions Evoked by Food Experiences: A Literature Review. *Frontiers in Psychology*, 9.
- Kenney E, Adhikari K. 2016. Recent developments in identifying and quantifying emotions during food consumption. *J Sci Food Agric*.
- King, S. C., Meiselman, H. (2010). Development of a method to measure consumer emotions associated with foods. *Food Quality and Preference*, 21, 168-177.

- King, S., Meiselman, H., Carr, B. T. (2010). Measuring emotions associated with foods in consumer testing. *Food Quality and Preference*, 21, 1114–1116
- Langlois, J., Dacremont, C., Peyron, D., Valentin, D., Dubois, D. (2011). Lexicon and types of discourse in wine expertise: The case of vin de garde. *Food Quality and Preference*, 22, 491-498.
- Laurans G., Desmet P. (2012). Introducing PrEmo2: new directions for the non-verbal measurement of emotion in design, in *Proceedings of the 8th International Conference on Design and Emotion*. London, United-Kingdom.
- Le Fur, Y., Teyssot, E., Vieira Foz, N., Foissy, C., Le Bras, G., Robbillard, B. (2010). Analyse lexicale autour des concepts de gras et volume dans les vins. *Revue Française d'œnologie*, 244, 20-30.
- Lerner, E.B., Jehle, D.V.K., Janicke, D.M., Moscati, R.M. (2000). Medical communication: Do our patients understand? *The American Journal of Emergency Medicine*, 18, 764-766.
- Leveau, N., Jhean-Larose, S., Denhière, G. (2011). EMOVAL : Evaluation automatique de la valence émotionnelle des textes à l'aide d'une métanorme émotionnelle de 5656 mots-racines. *Psychologie française*, 56, 209-221.
- Loewenstein, G., Friedman, Y., McGill, B., Ahmad, S., Linck, S., Sinkula, S., Beshears, J., Choi, J., Kolstad, J., Laibson, D., Madrian, B., List, J., Volpp, K. (2013). Consumers' misunderstanding of health insurance. *Journal of Health Economics*, 32, 850-862.
- Mauss, I., Robinson, M. (2009). Measures of emotion: A review. *Cognition & emotion*, 23, 209-237.
- Messina, D. Morais, J., Cantraine, F. (1989). Valeur affective de 904 mots de la langue française. *Cahiers de Psychologie cognitive*, 9, 165-187.
- Mojet, J., Dürrschmid, K., Danner, L., Jöchl, M., Heiniö R., Holthuysen, N. (2015). Are implicit emotion measurements evoked by food unrelated to liking? *Food Res. Int*, 76, 224–232.
- Monnier, C., Syssau, A. (2014). Affective norms for French words (FAN). *Behav. Res. Methods*, 46, 1128–1137.
- Nestrud, M., Meiselman, H., King, S., Leshner, L., Cardello, A. (2016). Development of EsSense25, a shorter version of the EsSense Profile. *Food Quality and Preference*, 48, 107–117.

- Ng, M., Chaya, C., Hort, J. (2013). Beyond liking: Comparing the measurement of emotional response using EsSense Profile and consumer defined check-all-that-apply methodologies. *Food Quality and Preference*, 28, 193–205.
- O'Connell, R.L., Hartridge-Lambert, S .K., Din, N., St John, E.R., Hitchins, C., Johnson, T. (2013). Patients' understanding of medical terminology used in the breast clinic. *The Breast*, 22, 836-838.
- Painchaud, L. (2005). *Le Halo Affectif des mots et leur répertoire*. Sherbrooke : Productions G.G.C.
- Parr, W.V., Mouret, M., Blackmore, S., Pelquest-Hunt, T., Urdapilleta, I. (2011). Representation of complexity in wine: Influence of expertise. *Food Quality and Preference*, 22, 647-660.
- Porcherot C., Delplanque S., Raviot-Derrien S., Le Calvé B., Chrea C., Gaudreau N. (2010). How do you feel when you smell this? Optimization of a verbal measurement of odor-elicited emotions. *Food Quality and Preference*, 21, 938–947.
- Rodrigues, S. (2016). "Minéralité des vins : parlons-en !" : la conceptualisation d'un descripteur sensoriel mal défini. *Alimentation et Nutrition*. Université de Bourgogne.
- Rogaux, M. (2010). Consumer understanding and reaction to health claims. Insights and methodology. In S. R. Jaeger & H. MacFie (Eds.), *Consumer-driven innovation in food and personal care products*. Cambridge: Woodhead. 277–302.
- Rose, G., Laing, D.G., Oram, N., Hutchinson, I. (2004). Sensory profiling by children aged 6–7 and 10–11 years. Part 1: a descriptor approach. *Food Quality and Preference*, 15, 585-596.
- Schutz, H.G. (1994). Appropriateness as a measure of the cognitive-contextual aspects of food acceptance. *Measurement of Food Preferences*, 25–50.
- Schouteten, J., De Steur, H., De Pelsmaeker, S., Lagast, S., De Bourdeaudhuij, I., Gellynck, X. (2015). An integrated method for the emotional conceptualization and sensory characterization of food products: the EmoSensory Wheel. *Food Res. Int.* 78, 96–107.
- Soares, A.P., Comesaña, M., Pinheiro, A.P., Simões, A., Frade, C.S. (2012). The adaptation of the Affective Norms for English words (ANEW) for European Portuguese. *Behavior Research Methods*, 44, 256–269.

- Spinelli, S., Masi, C., Zoboli, G., Prescott, J., Monteleone, E. (2015). Emotional responses to branded and unbranded foods. *Food Quality and Preference*, 42, 1–11.
- Stancu, V. (2006). Consumer understanding of health claims: methods for assessment and potential antecedents. PhD dissertation. Aarhus University. 226p.
- Stone, P. J., Dunphy, D. C., & Smith, M. S. (1966). *The general inquirer: A computer approach to content analysis*. M.I.T. Press.
- Syssau, A., & Font, N. (2003). Evaluation des caractéristiques émotionnelles d'un corpus de 604 mots. *Bulletin de psychologie*, 58, 361-367.
- Turney P. and Littman M. (2002). Unsupervised learning of semantic orientation from a hundred-billion-word corpus. Technical Report, National Research Council Canada.
- Turney P. and Littman M. (2003). Measuring Praise and Criticism: Inference of Semantic Orientation from Association. *ACM Transactions on Information Systems (TOIS)*, 21, 315-346
- Varela, P.A., Salvador, A., Gámbaro, A., Fiszman, S. (2007). Texture concepts for consumers: a better understanding of crispy-crunchy sensory perception. *Eur Food Res Technol*, 226, 1081-1090.
- Vikis-Freibergs, V. (1976). Abstractness and emotionality values for 398 French words. *Canadian Journal of Psychology*, 30, 22-30.
- Vincze, N., Bestgen, Y. (2011). Identification de mots germes pour la construction d'un lexique de valence au moyen d'une procédure supervisée. *TALN 2011*, 223–234. Montpellier, France.
- Waehrens, S., Zhang, S., Hedelund, P., Petersen, M., Byrne, D. (2016). Application of the fast sensory method 'Rate-All-That-Apply' in chocolate Quality Control compared with DHS -GC -MS. *International Journal of Food Science & Technology*, 51.
- Watson, D., Clark, L., Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* 54, 1063–1070.
- Worch, T., Lê, S., Punter, P. (2010). How reliable are the consumers? Comparison of sensory profiles from consumers and experts. *Food Quality and Preference*, 21, 309-2018.

