Understanding consumers' emotions and sensory experience for beauty care products

by

Congcong Zhang

B.S., Beijing Forestry University, 2014
M.S., The Ohio State University, 2016

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Food, Nutrition, Dietetics and Health
College of Health and Human Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Abstract

The beauty care space encompasses a diverse variety of product categories such as skin care, hair care, makeup, and others. In a fast-moving industry flooded with new products every year, cosmetic companies need to constantly offer new products that are adapted to the need and preferences of consumers to stay competitive in the marketplace. Understanding consumer experience related to hedonic, sensory, and emotional aspects of products is the key to driving consumer-centric product design for the beauty care category. This dissertation conducted three independent studies aiming to explore consumer experience of beauty care products from two perspectives: liking and beyond liking (emotions), based on conventional sensory and consumer data and online product reviews.

The objective of Chapter 2 was to develop an emotion lexicon that could be used to profile consumers' emotional responses to beauty care products in sensory and consumer tests. The lexicon was developed in four main steps: sourcing terms from online product reviews, term identification and categorization, term refinement, and term validation. The last three steps of the process were conducted with beauty care consumers using online surveys and interviews. The final emotion lexicon consists of 37 positive emotions and 20 negative emotions. The lexicon was able to discriminate product categories in testing product concepts. Recommendations on the application of this lexicon to each of the three categories of beauty care (skincare, hair care and makeup) were provided. The validated emotion lexicon from this study is readily applicable to other emotion research for skincare, hair care and makeup.

Chapter 3 explored sensory drivers of liking and emotional associations for beauty care products. Hand creams were used as testing samples to be evaluated for sensory characteristics and consumer perception. First, the sensory space (aroma, appearance, texture & skinfeel) of

twelve hand creams was profiled by a highly trained descriptive panel using a modified flavor/texture profile approach. Then, seven hand creams selected from the descriptive sensory space were rated for overall liking, emotions using the lexicon developed from Chapter 2, and consumer characterization using check-all-that-apply (CATA) in a home use test (HUT) with a hundred female consumers from the Kansas City area. Cluster analysis and external preference mapping identified three consumer clusters with different liking patterns: the thick & waxy-texture likers, mild scent & low-medium-thickness likers, and strong-scent likers. Consumers with different liking patterns differed in their emotional associations with sensory characteristics of hand creams. However, high intensities of certain aroma attributes seemed to elicit high-arousal emotions for all groups. Comparing consumer characterization using CATA with descriptive analysis configuration helped interpret these differences between consumer segments. In general, the differences in emotion associations across consumer clusters seems to be due to the values and efficacy each cluster attached to the sensory characteristics of the hand creams. The findings of this study could guide the development of new hand cream products targeting different consumer segments.

Chapter 4 explored consumer experience for hand cream products from the “voice of consumers”-online product reviews. A total of 17, 581 reviews representing 46 hand creams of different brands, price points, and sensory attributes were collected from Amazon and Ulta Beauty using a scraping software. Text analysis including text pre-processing, tokenization, frequency calculation, topic modeling and sentimental analysis were performed primarily using the tidytext package in R. Topic modeling using Latent Dirichlet allocation (LDA) identified five major topics consumers mentioned in these online reviews: greasiness & residue of the product, scent/fragrances of the product, skin feel & efficacy of the product, consumers’ skin issues, and occasions when to apply the product. In addition, term frequency–inverse document frequency (tf-idf) was calculated for each rating group to identify the most relevant unigrams and bigrams. It was found unpleasant scent and overall dissatisfied quality such as counterfeit product were the main reasons why consumers gave a rating lower than 4 stars. High efficacy such as moisturizing, and healing cracked skin, not greasy and smooth/soft skin feel were the drivers for a 5-star rating. These findings highlighted the importance of sensory experience and perception of efficacy in consumers' whole product experience. Consumer terminology regarding aroma, texture and skin feel of hand creams were also collected in this study.

Understanding consumers' emotions and sensory experience for beauty care products

by

Congcong Zhang

B.S., Beijing Forestry University, 2014
M.S., The Ohio State University, 2016

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Food, Nutrition, Dietetics and Health
College of Health and Human Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Approved by:

Major Professor
Martin Talavera

Copyright

Abstract

The beauty care space encompasses a diverse variety of product categories such as skin care, hair care, makeup, and others. In a fast-moving industry flooded with new products every year, cosmetic companies need to constantly offer new products that are adapted to the need and preferences of consumers to stay competitive in the marketplace. Understanding consumer experience related to hedonic, sensory, and emotional aspects of products is the key to driving consumer-centric product design for the beauty care category. This dissertation conducted three independent studies aiming to explore consumer experience of beauty care products from two perspectives: liking and beyond liking (emotions), based on conventional sensory and consumer data and online product reviews.

The objective of Chapter 2 was to develop an emotion lexicon that could be used to profile consumers' emotional responses to beauty care products in sensory and consumer tests. The lexicon was developed in four main steps: sourcing terms from online product reviews, term identification and categorization, term refinement, and term validation. The last three steps of the process were conducted with beauty care consumers using online surveys and interviews. The final emotion lexicon consists of 37 positive emotions and 20 negative emotions. The lexicon was able to discriminate product categories in testing product concepts. Recommendations on the application of this lexicon to each of the three categories of beauty care (skincare, hair care and makeup) were provided. The validated emotion lexicon from this study is readily applicable to other emotion research for skincare, hair care and makeup.

Chapter 3 explored sensory drivers of liking and emotional associations for beauty care products. Hand creams were used as testing samples to be evaluated for sensory characteristics and consumer perception. First, the sensory space (aroma, appearance, texture & skinfeel) of

twelve hand creams was profiled by a highly trained descriptive panel using a modified flavor/texture profile approach. Then, seven hand creams selected from the descriptive sensory space were rated for overall liking, emotions using the lexicon developed from Chapter 2, and consumer characterization using check-all-that-apply (CATA) in a home use test (HUT) with a hundred female consumers from the Kansas City area. Cluster analysis and external preference mapping identified three consumer clusters with different liking patterns: the thick & waxy-texture likers, mild scent & low-medium-thickness likers, and strong-scent likers. Consumers with different liking patterns differed in their emotional associations with sensory characteristics of hand creams. However, high intensities of certain aroma attributes seemed to elicit high-arousal emotions for all groups. Comparing consumer characterization using CATA with descriptive analysis configuration helped interpret these differences between consumer segments. In general, the differences in emotion associations across consumer clusters seems to be due to the values and efficacy each cluster attached to the sensory characteristics of the hand creams. The findings of this study could guide the development of new hand cream products targeting different consumer segments.

Chapter 4 explored consumer experience for hand cream products from the "voice of consumers"-online product reviews. A total of 17, 581 reviews representing 46 hand creams of different brands, price points, and sensory attributes were collected from Amazon and Ulta Beauty using a scraping software. Text analysis including text pre-processing, tokenization, frequency calculation, topic modeling and sentimental analysis were performed primarily using the tidytext package in R. Topic modeling using Latent Dirichlet allocation (LDA) identified five major topics consumers mentioned in these online reviews: greasiness & residue of the product, scent/fragrances of the product, skin feel & efficacy of the product, consumers' skin issues, and occasions when to apply the product. In addition, term frequency–inverse document frequency (tf-idf) was calculated for each rating group to identify the most relevant unigrams and bigrams. It was found unpleasant scent and overall dissatisfied quality such as counterfeit product were the main reasons why consumers gave a rating lower than 4 stars. High efficacy such as moisturizing, and healing cracked skin, not greasy and smooth/soft skin feel were the drivers for a 5-star rating. These findings highlighted the importance of sensory experience and perception of efficacy in consumers' whole product experience. Consumer terminology regarding aroma, texture and skin feel of hand creams were also collected in this study.

Table of Contents

List of Figures

List of Tables

Acknowledgements

The past three years have been the most joyful and challenging journey in my life. I would not have completed this dissertation without the support of my professors, graduate friends, co-workers, and parents. First, I would like to thank my major advisor Dr. Martin Talavera for his guidance and support throughout my PhD study. He was such an open-minded mentor who kept inspiring me to ask valuable research questions and encouraged me to work for the answers. I would also like to thank Dr. Edgar Chambers IV, Dr. Marianne Swaney-Stueve, Dr. Michael Higgins, and Dr. Wendi Armbrister for providing insightful and constructive advice on this research as members of my dissertation committee. I'm so grateful to Dr. Marianne Swaney-Stueve and Dr. Janelle Elmore for being a great source of inspiration for me when designing consumer studies.

I would like to thank Bert, Tyler, Jessie, and Bethany for their help with my projective mapping study. I'm thankful to Gongshun, Porranee and Gagan who helped me with descriptive analysis in Manhattan. I feel lucky to have Chetan, Grace, Huizi, and Denis around who were always there to answer my millions of technical questions. Thanks to Nancy, Kent, Barb, and Kay for their kindness and all the funny stories we talked while working in the center.

Finally, I would like to acknowledge my parents for their support and everything they did for me.

Chapter 1 - Literature Review

Emotion Measurement in Sensory and Consumer Research

Theoretical Ground and Definition of Emotion

The study of human emotion dates to the Socrates' time (470-399 BC) when philosophers have started discussing this topic and laid the foundation for many contemporary research traditions (Coppin & Sander, 2016). Since the 1980s with the advancement in experimental psychology and neuroscience, there has been an explosion of scientific research focused on conceptualizing and measuring emotions (Coppin & Sander, 2016). Many theories have been developed with attempts to describe what emotion is, how emotion is elicited and how emotional responses are organized. Yet, there hasn't been a consensus on the definition, content, and structure of emotions among scholars. Coppin & Sander (2016) have grouped the theories of emotions in three major families: basic emotion theories, dimensional theories, and appraisal theories. A consensual view reflected from these theories is that emotion is a fast and event-focused phenomenon with multiple components (expression, action tendency, bodily reaction, feeling and appraisal) (Coppin & Sander, 2016). Different theories tend to emphasize different components of emotions (however, they can overlap) and are related to specific measures.

Basic Emotion Theories

Rooted in evolutionary psychology, the central idea of basic emotion theories is that emotions enable the individual to handle fundamental life tasks—responding adaptively to threats and opportunities in the environment (Ekman, 1992). Basic emotion theorists assume there is a limited number of emotions that are biologically and psychologically "basic" for human beings (Wilson-Mendenhall et al., 2013). By contrast, complex emotions were thought to be the result of a mixture of basic emotions by some researchers (Tomkins, 1963; Ekman, 1992).

Others considered complex emotion phenomenon to include complex appraisals or higher order cognition such as thought and judgement that differ across individual or culture (Izard, 2007). While many researchers advocate the theories of basic emotions, there is continuous disagreement on the exact number of emotions that are qualified as basic emotions. For instance, Ekman proposed six basic emotions as fear, anger, sadness, disgust, surprise, and joy, which are widely known as "the Big Six" (Ekman et al., 1969). Later, he expanded this list by including amusement, contempt, contentment, embarrassment, excitement, guilt, pride in achievement, relief, satisfaction, sensory pleasure, and shame (Ekman, 1999). Izard (1979) offered 10 basic emotions: fear, disgust, anger, distress-anguish, surprise, interest-excitement, joy, contempt, shame, and guilt. Plutchik, (1980) proposed eight primary emotions: joy, sadness, acceptance, surprise disgust, fear, anger, and anticipation which were arranged in a color wheel.

As basic emotions are believed by basic emotion theorists to be innate and have distinctive universal expression which can be universally recognized, measurement of such emotions have been focused on measuring the physiological associations of emotions such as facial expressions which are considered as an involuntary motor manifestation of the autonomic nervous system activity Ekman (1992). However, this approach has been criticized. For example, Russell (1994) pointed out that the recognition of facial expressions of emotion are not universal as previously thought.

Dimensional Theories

Contrary to basic emotion theories which assume emotions are characterized as discrete entities, dimensional theories suggest that emotions are fundamentally the same and any emotional states can be projected to the affective space consisted of several dimensions (Bestelmeyer et al., 2017). Dimensional theories focus on the conscious subjective affective experience components of emotions, which are generally referred to as feelings. Several models have been developed to conceptualize emotions in the affective dimensions (Coppin & Sander, 2016). Of these models, Russel's circumplex model of emotion which characterizes emotional states in a bidimensional space has been one of the most influential. In this model, all emotions are constructed from two independent dimensions which are valence (negative/unpleasant-positive/pleasant) and arousal (sleepy/deactivation-activation) (Posner et al., 2005; Russell & Barrett, 1999; Russell & Pratt, 1980). Accordingly, emotional states located in different areas of the circumplex represent various levels of valence and arousal (Figure 1.1) (Posner et al., 2005). Based on this circumplex model, Russel introduced a single-item scale—the Affect Grid to quickly measure affect across the valence-arousal space (Russell et al., 1989). On the Affect Grid that contains 9×9 squares with anchor emotion words in the corners, respondents can select one of the 81 squares to indicate their feelings (Figure 1.1) (Russell et al., 1989).

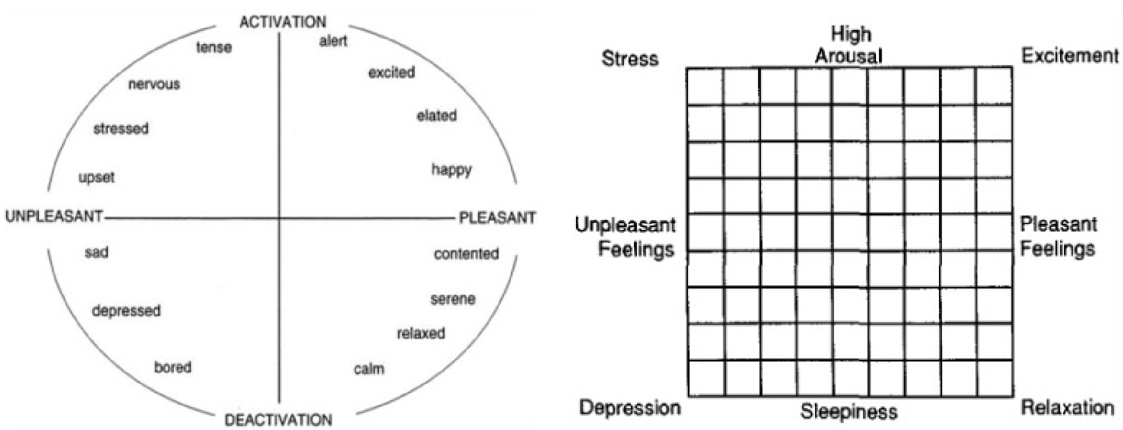


Figure 1.1. The circumplex model of affect (left) and the Affect Grid (right) to measure affect across balance-arousal space

Another influential bidimensional model is the Positive Affect-Negative Affect proposed by Watson and Tellegen (1985). This model was constructed on two orthogonal bipolar dimensions of mood: high positive affect-low positive affect and high negative affect-low negative affect (Watson & Tellegen, 1985). Based on the work of Zevon & Tellegen (1982) and

Watson & Tellegen (1985), the Positive and Negative Affect Schedule (PANAS), a self-reported affective measurement tool, was developed to assess feelings and moods (Watson et al., 1988). The PANAS consists of two scales with 10 items for the Positive Affect (PA) scale (active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud, strong) and 10 for the Negative Affect (NA) scale (afraid, scared, nervous, jittery, irritable, hostile, guilty, ashamed, upset, and distressed). Each item is rated on a 5-point scale ranging from "very slightly or not at all" to "extremely." The scores of the two-primary dimensions of mood (PA and NA) are sums of the ratings of the PA and NA items. Since the publication of the PANAS, it has become one of the most widely used affective measures in psychology and other disciplines such as health-behavior research and sensory and consumer research. For instance, Kuesten et al. (2014) conducted a global study that used the PANAS scales to measure consumer emotions associated with phytonutrient supplements of different aromas. The results of this study showed the PA and NA were able to differentiate products as well as users of products, suggesting the PANAS as a useful tool in the measurement of consumers' emotional responses to products (Kuesten et al., 2014). Despite the popularity of the bidimensional emotion models in measuring subjective emotional experience, the two-dimensional construct has been criticized such as being unable to represent the whole world of emotions (Fontaine et al., 2007) and lacking important verbal descriptions to characterize feelings induced by specific stimuli in practical field such as consumer research (Chrea et al., 2008).

An Integrative Framework- Hierarchical Model of Emotions

Basic emotions are narrow and specific, while dimensional theories conceptualize emotions in general dimensions such as positive and negative affect. This divergency between the two levels of understanding has been integrated within hierarchical structures. In a typical hierarchical framework proposed by Shaver et al. (1987) and Storm and Storm (1987), the superordinate consists of positive and negative affect, followed by the next level as the basic emotion levels, and the lowest level (subordinate) is made up of groups of individual emotions that are classified as the most typical emotions of each basic emotion category (Shaver et al., 1987; Storm and Storm, 1987). Following the principles of the hierarchical structure of emotions, a hierarchical of emotions associated with food was created by Laros and Steenkamp (2005) with a superordinate level labeled as positive affect and negative affect, second level with four selected basic emotions (love and pride were excluded) of each positive and negative affect, subordinate level with specific emotions obtained from the Consumption Emotion Set (Richins, 1997). The emotion terms proposed in this model were rated across different types of food (genetically modified food, functional food, organic food, or regular food) in a consumer survey. The results of the test supported the proposed 3-level hierarchical model and also suggested basic emotions provide more information for better understanding consumer feelings related to food compared to only positive and negative affect. However, the authors recommended that to use either part of the model or the whole model should depend on the research question (Laros & Steenkamp, 2005).

Appraisal Theories

Another major theory of emotions are the appraisal theories, which focus on the process of emotion elicitation. Appraisal theorists believe that the elicitation of emotions is caused by the evaluation of a stimuli or situation based on different criteria such as novelty, relevance, intrinsic pleasantness (or unpleasantness), predictability, goal-relevance, coping potential, and normative significance (Coppin & Sander, 2016). Lazarus (1991) emphasized that it was the evaluation, rather than the stimulus itself that caused the emotion (Lazarus, 1991). Different from basic emotion theories, appraisal theorists assume that there are infinite number of emotions without considering some are more basic than others (Coppin & Sander, 2016). Also contrary to dimensional theories, appraisal theories consider each emotion has a distinctive appraisal process so that they cannot be combined in broad factors/dimensions (Frijda et al., 1989; Smith & Lazarus, 1993). Coppin & Sander (2016) summarized the ways that have been proposed for the measurement of appraisals including questionnaires, facial expressions, voice, psychophysiological responses, and brain imaging, while there has been no consensus on the specific way to measure it.

Appraisal theories have contributed to the interpretation of emotions related to consumption or sensory experiences. For example, (Desmet & Schifferstein, 2008) used appraisal theories to differentiate positive and negative consumption emotions from perspectives of the process that elicited emotions: positive emotions were evoked by events appraised as matching with personal concerns, while negative emotions were elicited by events that were appraised as in conflict with personal concerns. In addition, positive emotion, such joy, hope, fascination, satisfaction, and contempt are essentially different in the conditions that elicit them, which could be explained by their differences in the appraisal types (Schifferstein & Desmet, 2010). Appraisal theories suggest the importance of understanding the evaluation process underlining different emotions in developing products with features that could elicit pre-defined emotions. Moreover, appraisal theories have been used to explain and predict the emotional effects of olfactory stimulation (Chrea et al., 2008). The context of product consumption has been emphasized by appraisal theories as an emotion is elicited not only by the evaluation of the stimulus, but also the evaluation of the social and situational settings in which the emotion is experienced (Barrett et al., 2007).

Emotions in Sensory and Consumer Research

In the past decade, there has been huge progress in considering other aspects of consumer behavior than conventional acceptance or preference in consumer product research. Product-elicited emotions are among these hot topics. The increasing interests in emotional responses to food, beverage and personal care products are driven by their potential to provide insights beyond conventional consumer hedonics which helps with better product differentiation and communication. Instead of trying to understand what an emotion is and its elicitation mechanisms as they have done in theoretical research, the studies of emotion in sensory and consumer area have mainly focused on understanding how products are differentiated in emotional responses/affective feelings, as well as how to measure these emotions without imposing a strong theoretical framework. For instance, emotion lexicon and questionnaires developed by consumer researchers include both basic emotions and secondary/complex emotions (Bhumiratana et al., 2014; Ferrarini et al., 2010; King & Meiselman, 2010) and have been constructed in multiple dimensions using data-driven approach without following the valence-arousal models (Chrea et al., 2008; Richins, 1997).

Theoretical context has been included in interpreting emotional responses and developing emotional terms, however, the definition of emotion seems to be less of a debate in the practical field of consumer product research (Thomson & Crocker, 2013). This is because sensory and consumer research mainly uses consumers' emotional responses to discriminate products or to add new information beyond sensory and hedonic measures (Jiang et al., 2014; King et al., 2013; Porcherot et al., 2010). Researchers differentiate different affective feelings by mentioning that emotions are rapid, intense, focused on a referent, moods are long-lasting, diffuse without reference, and attitudes are more evaluative (Ferrarini et al., 2010; King & Meiselman, 2010).

However, emotions and feelings are often used interchangeably in practical research, and a broader scope of emotion has been considered as subjective affective feelings which may include emotions, moods, attitudes when selecting emotion terms for lexicon/questionnaire development (Chrea et al., 2008; Ferdenzi et al., 2013, 2013; King & Meiselman, 2010; Spinelli et al., 2014; Thomson & Crocker, 2013). Researchers also included abstract conceptualizations with emotional connotations in emotional questionnaires (Thomson et al., 2010). This broad sense was explained by Spinelli et al. (2014) who stated that both emotions and other affective feelings were of the interest in understanding consumer experience as product perception was mediated by not only emotions elicited at the moment of consumption, but also by feelings associated with products in the mind of consumers.

Emotion Research for Beauty Care

The study of emotions elicited by personal care products has been of interest to researchers in the sensory and consumer field. This is driven by its potential to provide insights in addition to traditional consumer liking and help with better product differentiation and communication. For example, in the evaluation of perfumes with 61 consumers, Porcherot et al. (2010) found that two perfumes of the same perceptual category receiving similar hedonic ratings differed in the feelings they induced. Moreover, understanding the emotional experiences that consumers seek and perceive from a product would help companies to design their product in a way that reinforces the desired values and benefits (Porcherot et al., 2010).

For the personal care category, product-elicited emotions are derived from the smell, touch, and visual experience of the product, though most of the published research has focused on odor-induced emotions. Different explicit and implicit measurement tools have been developed and applied to measure odor-elicited emotions. The Geneva Odor Scales (GEOS) developed by Chrea et al. (2008) in Switzerland has been a pioneer method that verbally measures feelings induced by fragrances and odors. The original GEOS questionnaire requires respondents to rate the intensities of 36 emotion terms in 6 dimensions including pleasant feeling, unpleasant feeling, sensuality, relaxation, refreshment, and sensory pleasure while smelling an odor (Chrea et al., 2008). To adapt the GEOS for a commercial setting, a shorter format was later developed; it included only the three most representative terms in each of the six dimensions and respondents rate the intensity of each emotional dimension (Porcherot et al., 2010). The simplified version of GEOS has shown comparable results to the complete version when tested in the evaluation of 12 shampoos (blind coded and neutral packaging) with 78 participants in sensory booths (Porcherot et al., 2010). Taking cultural differences into account, the GEOS has been extended to a series of culture-specific Emotion and Odor Scales (EOSs) by applying the same development procedure as used in the Swiss culture to 5 other cultures (Brazil, China, Singapore, United Kingdom, and the United States) (Chrea et al., 2008). Combining the common categories of emotions and culture-specific categories from the EOSs, the Universal Emotion and Odor Scale (UniGEOS) was then developed to meet the needs for multi-country consumer research (Ferdenzi et al., 2013). The creation of the EOS and UniGEOS illustrated systematic approaches in developing explicit verbal tools that can be used for the measurement of emotions elicited by odors and fragrances in multi-culture consumer research settings.

In addition to verbal self-reported questionnaires, Churchill & Behan (2010) have applied a visual approach – the Mood Portraits test – to get self-reported emotional responses using a series of photographs of scenes that evoke different emotional responses. In this test, respondents were asked to select pictures that evoked a similar mood comparable to each of the 10 fragrance samples (Churchill & Behan, 2010). This type of non-verbal measurement technique has the advantages of being language independent and more importantly requiring less cognitive interpretation, thus capturing more spontaneous emotional responses from the respondents (Churchill & Behan, 2010).

As an alternative to explicit measurement, implicit techniques such as physiological measures, facial expression measures, and behavioral measures have gained popularity in the consumer research field as emotional responses are collected in an indirect and non-self-reported way (Lagast et al., 2017). For example, Painchault et al. (2020) demonstrated the relaxing effect of the Peony fragrance in hair care products by examining the physiological, mental and physical, and behavioral responses of 55 women to scented shampoos and hair serums. Physiological measures such as heart rate, blood pressure, and electrodermal activity were cited as the most effective measures for demonstrating relaxation; the results from these measures were significant enough to support a product claim on the "emotional benefits" of the aroma (Painchault et al., 2020). This represents a novel approach to claims testing and demonstrates that implicit measures can be effective in characterizing consumer's emotional responses to a product based on its fragrance. Despite the advantage of being not under the conscious control of the respondents, implicit techniques such as physiological measures have been shown unable to differentiate a large number of different positive emotions (Meiselman, 2016). Combining implicit measurement with explicit measurement may provide insights to better understand emotions from both conscious and unconscious levels. For example, in a study looking at product effects on emotional states, David et al. (2019) provided participants (N=26) with an odorless and rose-scented cosmetic cream and asked them to complete a survey on their emotional state ("mood") and undergo face analysis and an MRI to record their brain activation patterns. More positive results were found for the scented cream from brain-related measures

(David et al., 2019). As supplementary information, more relaxation, satisfaction, happiness, and less stress were found for the scented cream if presented after the odorless cream from the self-reported survey (David et al., 2019). The authors acknowledged the limitations of the small sample size and indicated more testing was needed (David et al., 2019). The lower sample size was also noted as being an issue for implicit measurement and implicit combined with explicit measurement in cases of emotion research associated with food products (Lagast et al., 2017).

Other than odor-elicited emotion research, emotions elicited by the tactile experiences of products have been studied as well. (Guest et al., 2011) developed the Tactile Perception Task (TPT), containing a set of sensory and emotional attributes that can be used for the assessment of emotions elicited by personal care products with a variety of textures. In addition, Kao Corporation, a Japanese chemical, and cosmetic company, has been conducting a series of studies exploring relationships between texture perception and emotions. As published in the company's news release, an emotional assessment scale containing 12 factors has been created for the evaluation of emotions associated with tactile stimulation (Kao Corporation, 2018a). It was also found that tactile stimulation during skin care can elicit various pleasant feelings (Kao Corporation, 2018a). Another study conducted by Kao Corporation indicated that skincare products of different textures could evoke different emotions during product application (Kao Corporation, 2018b). In this study, two groups of women (N=29 for each group) were recruited to apply each of four fragrance-free and color-free creams with different textures for 3 consecutive days (Kao Corporation, 2018b). After the three days, participants evaluated the texture of each sample and reported their emotions during application of night skin care based on the emotional assessment scale (Kao Corporation, 2018b). Correspondence analysis of texture results and emotional responses showed that creams in different quadrants on the texture map corresponded to the locations on the emotional map, indicating that different textures can elicit different emotions (Kao Corporation, 2018b). Although these lacked detailed study design or comparison of the results to peer review publications, they demonstrated the need for investigating texture-elicited emotions and discussed methods for conducting this type of research in a home-use setting.

It is important to note that emotions evoked by a personal care product might not only be due to one or two sensory modalities of the product, but also the multisensory experience, expectations, product functions, and extrinsic characteristics such as the brand and packaging. Talavera & Sasse (2019) gathered positive and negative emotional terminology for each of the makeup, skincare, and haircare categories by conducting focus group discussions with beauty care consumers. The emotional responses collected are based on the whole product experience instead of a single perceptual modality, which would be useful for understanding product experience and product communication with consumers. Such terminology can be used as the starting point to develop product-specific emotion lexicons for different personal care categories.

Sensory and Consumer Research on Beauty Care Products

The beauty care space encompasses a diverse variety of product categories such as skin care, hair care, makeup, fine fragrance, and others. In a fast-moving industry flooded with new products every year, cosmetic companies need to constantly offer new products that are adapted to the need and preferences of consumers in order to stay competitive in the marketplace. The potential success of beauty care products depends not only on their efficacy, but also the sensory characteristics, consumer hedonic and emotional experiences (Parente et al., 2011; Wortel & Wiechers, 2000). Consequently, sensory and consumer research is essential in driving the product development of beauty care. Historically developed for the food and beverage industry, sensory and consumer research techniques have been now widely applied to the evaluation of beauty care products to provide insights guiding product innovation and maintenance. The following section gives a summary of the adaptation of sensory and consumer research methods in the beauty care category.

Consumer Research

In consumer studies of beauty care products, qualitative methods such as Focus Groups and In-depth Interviews have been conducted to gather consumer terminology and gain insights into consumer expectations, attitudes and their rationales behind their choices (Kahraman & Kazançoğlu, 2019; Talavera & Sasse, 2019). Special attention must be paid if the interest of the study is in the application or the afterfeel of the products. In this case, participants might be asked to first test the product at home, take notes or pictures before coming to a focus group or interview (Dreyfuss, 2018). Quantitative consumer research approaches such as central location test (CLT) and home-use-test (HUT) have been used to quantify product appeal, and consumers' sensory and emotional experiences (Bourguet et al., 2016; Parente et al., 2011).

Due to the distinctive category characteristics, multiple elements specific to beauty care products should be considered when testing these products with consumers in comparison to testing food and beverages (Delarue et al., 2018). First, because of the diversity of product categories, usage and functionalities of beauty care, consumer tests must consider several aspects of product experiences including sensory experiences, comfort in usage and product functionality depending on the type of beauty product under testing (Delarue et al., 2018). In specific, product performance such as efficacy is the predominant part to be evaluated to understand consumer hedonics and perceived benefits that can be used to support claim substantiation (Delarue et al., 2018; Dreyfuss, 2018). The efficacy of beauty care products is inseparable from sensory experience, for example, a color that is not intense or a texture that is too thin for a product might result in the perception of a lack of efficacy (Dreyfuss, 2018). It is usually reflected from the sensations perceived after product application. To understand and claim long-term efficacy of some products such as fragrances, lipsticks, or hairsprays, follow-up interviews or surveys with consumers about the continuing effects or "hold" of a product might be needed (Dreyfuss, 2018). In addition, compared to food and beverages, more practical constrains might be considered when testing with beauty care products such as difficulties in blinding the samples and environmental control, especially in the case of CLT. For instance, testing with skin care products might require the test facilities equipped with a mirror, a washbasin, temperature control, uniform lightening system, and an air renewal system (Delarue et al., 2018; Martins et al., 2020; Parente et al., 2011). Moreover, for the beauty care category, it is important for the product to be used and evaluated for its intended purpose in its natural usage context so that consumers experience different ways of interacting with the products (Delarue et al., 2018). Because of this, it is more frequent for beauty care products to be tested in HUT compared to CLT. Finally, when recruiting participants for beauty care studies, individual differences, and segmentations such as skin/hair types, habits, routines should be considered, depending on the product type and objectives of the studies (Delarue et al., 2018). Participants in HUT are usually asked to suspend the use of their current product.

Sensory Drivers of Liking and Emotions

At early stages of beauty product development, it is key for product developers to understand which sensory attributes drive consumer liking, choice, and perceived efficacy of the product category. Studies have been published that employed consumer data to uncover sensory drivers of liking for beauty care products. For instance, to understand the drivers of liking for antiaging creams, Parente et al. (2011) proposed an external preference mapping approach that was based on consumers' hedonics and responses to a Check-All-That-Apply (CATA) question. In a central location test, sixty-nine consumers were asked to rate their overall liking and answer a CATA question consisting of a list of 42 terms towards six anti-aging creams (Parente et al., 2011). The 42 CATA terms were selected from previous studies, published data as well as market campaigns, which represented different aspects of product experience: sensory characteristics, emotional associations, effects on the skin, product positioning, and applications (Parente et al., 2011). The sample configuration was derived from multiple factor analysis (MFA) of CATA count, considering different categories of CATA terms as separate groups and overall liking scores as supplement variable (Parente et al., 2011). External preference mapping was used to regress consumer overall liking against the samples' coordinates from the first two dimensions of MFA in quadratic response surface models (Parente et al., 2011). This analysis resulted in a density plot that suggested the area and direction of maximizing consumer overall liking. This study indicated the sensory characteristics that were the main determinants for consumers' hedonic and emotional responses to different antiaging creams (Parente et al., 2011). Products' hydrating ability, texture and perfume were found to be the main drivers of liking for antiaging creams (Parente et al., 2011). The most liked products elicited more positive emotions among consumers (Parente et al., 2011). Utilizing consumer CATA responses for product profiling, this study demonstrated a quick and simple alternative to conventional external preference mapping that uses sensory data collected from trained panel.

In addition, projective mapping coupled with CATA, flash profile and intensity scales have been cited for sensory characterization of skin care products using untrained panel/consumers (Martins et al., 2020; Parente et al., 2010, 2015). These rapid methods provide not only more flexibility and time/cost efficiency in the product development of beauty care, but also valuable information about consumers' perception of a product category that can be used in product positioning and advertisement. Despite these advantages, they cannot replace conventional descriptive analysis that use trained panels in various cases, for example, when subtle differences among products are important (or when testing products are very similar), and when comparing sensory profiles of products over time (Varela & Ares, 2012). Moreover, most rapid methods commonly require simultaneous presentation of all products for a direct comparison, which may be difficult to achieve in various cases of beauty product evaluation, such as testing the long-lasting effects of cosmetic products on skin, or the performance of hair care products on hair (Dreyfuss, 2018).

Besides preference mapping, researchers have used design of experiment (DOE) to identify the optimum combination of product characteristics that maximizes consumer acceptance of beauty care. Aiming to explore the best level combinations of formula, thickness, weight, and lotion add-on for makeup remover wipes, Xing et al. (2020) conducted a 6-week HUT with 963 consumers in the United States and United Kingdom. 18 prototypes with different combinations of four physical factors were generated using a quasi-central component design (Xing et al., 2020). The HUT followed a balanced incomplete block design (BIBD) with each participant evaluating 6 of the 18 prototypes (Xing et al., 2020). In each week, participants were asked to use each prototype at least once a day for 5 days to remove makeup and complete an online survey at the end of the week (Xing et al., 2020). Overall liking (OL), purchase intent (PI), likelihood to recommend to a friend, new and different, and product performance related questions such as durability, texture of the fabric, ease of use, and skin feel were asked in the online survey (Xing et al., 2020). Optimization modeling and sensitivity analysis were performed to understand which DOE factors had higher impact on consumer OL and PI (Xing et al., 2020). At the same time, driver's analysis and penalty analysis were used to uncover the key perceived attributes driving consumer OL and PI (Xing et al., 2020). With multiple analysis methods, this study provided a comprehensive understanding of the relative importance of different product features from both DOE and consumer perception perspectives (Xing et al., 2020).

Overall, there were only limited publications relevant in consumer research of beauty care products. This type of research is normally conducted by cosmetic companies who tend to keep their studies confidential. Meanwhile, the only few drivers of liking studies published are all based on consumer hedonics and perception. Future research could utilize conventional external preference mapping techniques based on consumer hedonics and sensory data from trained panel to validate the findings of previous research. Preference mapping techniques can also be adapted to explore drivers of benefits or perceived efficacy for beauty care products, a "hydrating map" for skin moisturizer for example (Dreyfuss, 2018).

Sensory Evaluation by Trained Panel

Consumers are probably able to differentiate between products, however, they might find it hard to describe the detailed attributes that differentiate two products from each other. Sensory evaluation by trained panel fills this gap by identifying and quantifying the key sensory characteristics of different products, which provides essential guidance for beauty product development (Meilgaard et al., 2016). In addition, results from well-designed and tightly controlled descriptive analysis following industry standards can be used to support claim substantiation for beauty and personal care products (ASTM Standard E1958 − 20, 2020; Churchill & Greenaway, 2018). As early as the 1970s, Schwartz (1975) from General Food has modified the sensory texture profile method that has been used in food evaluation in order to accommodate the special needs involved in the assessment of skincare products. In that modification, three-stage evaluation (pick-up, rub-out, and afterfeel) was proposed for the evaluation of skin care products (Schwartz, 1975). The multi-stage evaluation process has been used frequently in later research such as in the evaluation of emollients and characterization of the performance of cosmetic ingredients and products (Parente et al., 2015; Wortel & Wiechers, 2000). To meet the needs of more standardized guidance to conduct sensory evaluation in the beauty care field, standards organizations, such as ASTM, have published a series of guides for the assessments of skin creams and lotions, shampoos, fragrance/odors for hair products, deodorants (ASTM Standard E1207 − 14, 2014; ASTM Standard E1490-19, 2019; ASTM Standard E1958 − 20, 2020; ASTM Standard E2049 − 20, 2020; ASTM Standard E2082 − 12, 2020). Sensory evaluation of beauty care products requires special considerations due to the need to apply products on panelists' skin, face, or hair to simulate real in-use conditions without invalidating good sensory practice. Strict and detailed procedures of product handling and manipulation should be given. This could include standardizing the way and amount of sample to be delivered to the testing area, specifying the time point to perform the specific manipulation or evaluation, applying equal degrees of movement and pressure, using the same fingers when manipulating the product, applying the product to the same skin area (ASTM Standard E1490-19, 2019; Churchill & Greenaway, 2018). In addition, the evaluation process is dynamic; language used normally covers product attributes and skin/hair feel across multiple stages of product application. Table 1.1 summarized the recent application of sensory evaluation of different beauty and personal care products. The evaluation methods and languages used depend on the product type and objectives of the study. As referenced from Table 1.1, skincare products (even though intended for face use) are normally tested on the inner side of panelists' forearms. This provides the opportunity to test more than two products at a time compared to using face skin as a testing area (Schwartz, 1975). In cases of other product category, artificial objects such as hair swatches or artificial nails have been used in the descriptive evaluation (ASTM Standard E2082 − 12, 2020; Sun et al., 2014). Before product application on skin or hair, the testing area is normally pre-conditioned following a wash/wipe procedure using mild, non-perfumed soap/wipes and a drying-out period (ASTM Standard E1490-19, 2019; Aust & Oddo, 1989; Dooley et al., 2009; Lee et al., 2005). Moreover, as temperature and humidity of both testing area and the finger used for manipulation may influence the sensitivity of the measurement, ASTM suggests recording these two parameters for all panelists (ASTM Standard E1490-19, 2019). Based on the product type being evaluated, specific skin type might be screened before evaluation (ASTM Standard E1490-19, 2019). Finally, typical environment control includes constant level of temperature, humidity, air flow, and lighting (ASTM Standard E1490-19, 2019).

Table 1.1. Recent application of sensory evaluation of different beauty and personal care products

Study type/objective	Product	Stages	Sensory modalities	Testing site	Testing Area Preparation	Reference
Lexicon development	Lotion and creams	Before application, pick up, rub out, 2 min after application, 10 min after application	Appearance, texture, and skin feel	Back of the hand	Cleansing foam wash	(Lee et al., 2005)
Sensory characterization	Emollient	During application, immediate, 5min, 10min after application	Appearance, texture, and skin feel	Internal side of the non-dominant forearm	A mixture of isopropyl alcohol and water (45:55)	(Parente et al., 2008)
Sensory characterization	Ingredients	Before rubbing, during rubbing, after rubbing	Appearance and skin feel	Both inside and outside of left and right forearm	N/A	(Wortel & Wiechers, 2000)
Sensory characterization	Skincare products	Before rubbing, during rubbing, after rubbing, 1 h after application	Appearance, skin feel and skin sensation	N/A	N/A	(Wortel & Wiechers, 2000)
Sensory characterization	Cosmetic creams	During application	Texture and skin feel	Lateral side of the dominant forearm	A mixture of isopropyl alcohol and water (45:55)	(Boinbaser et al., 2015)
Lexicon development	Lip products	Before application, immediate after application, 10 min after application	Appearance and texture	Inside forearm and paper (color evaluation)	Fragrance-free, alcohol-free wipes	(Dooley et al., 2009)
Sensory characterization	Toothpaste	During application, after application (rinsing)	Flavor and texture	Inside mouth, teeth	Water and crackers	(Hightower & Chambers, 2009)

Study type/objective	Product	Stages	Sensory modalities	Testing site	Testing Area Preparation	Reference
Lexicon development	Cosmetic powders	Before application, pick up, during application, after application	Appearance, texture, and skin feel	Top of the hand	N/A	(Moussour et al., 2017)
Sensory characterization	Emulsion	Before application (in a jar), pick up, during application, after application	Appearance, texture, and skin feel	Top of the hand	N/A	(Moussour et al., 2017)
Sensory characterization	Hair	Before manipulation, during manipulation, after manipulation	Appearance and texture	Hair samples	Pre-wash with sodium lauryl ether sulfate 10% w/w for 1 min, rinsed for the same time with flowing water	(Bloch et al., 2021)
Sensory characterization	Toothpaste	During tasting	Flavor and mouthfeel	Inside month	Warm water, whole milk, rice cake stick	(Kim et al., 2013)
Lexicon development	Nail polish	Application and removal	Appearance, aroma, and texture	Artificial hands and nails	N/A, acetone used for removal	(Sun et al., 2014)

Most of the sensory literature relevant to the beauty care category focuses on skin creams or lotions; the purposes of these studies are either to develop a lexicon for a category of skin care products (Lee et al., 2005) or to characterize the performances of different formulations or cosmetic ingredients (Boinbaser et al., 2015; Parente et al., 2008; Wortel & Wiechers, 2000). A descriptive lexicon consisted of 26 attributes with corresponding definitions and references has been developed by a group of 10 trained panelists for the evaluation of aqua creams marketed in Korea (Lee et al., 2005). This lexicon included attributes in five stages: appearance (viscosity, gloss, transparency), pick-up (firmness, amount of peaking, cohesiveness), rub-out (spreadability, silkiness, thickness, coolness, wetness, oiliness, adhesiveness, absorbency), after-feel with 2 minutes and 10 minutes after application (gloss, stickiness, oiliness, moisturization, silkiness, amount of residue) (Lee et al., 2005). A perceptual map was then constructed within the space of 12 aqua creams, which indicated that these aqua creams were mainly differentiated on their oiliness, viscosity, adhesiveness, and thickness in contrast to transparency, wetness, coolness and spreadability (first dimension of principal component analysis), as well as stickiness and gloss (second dimension of principal component analysis) (Lee et al., 2005). Spreadability (rub-out), oiliness and stickiness (afterfeel), gloss (appearance and afterfeel), and amount of residue have also been the main attributes in sensory characterization of emollients (Parente et al., 2008) and other personal care ingredients (Wortel & Wiechers, 2000). All these attributes can be adapted for the evaluation of other skincare products such as hand creams and lotions.

Sensory profiles of beauty care products changes overtime during and after application due to the changes in their rheological properties, water evaporation and absorption into the stratum corneum of the skin (Boinbaser et al., 2015). To capture these dynamics, Boinbaser et al.

(2015) proposed a novel temporal methodology, Temporal Check-All-That-Apply (TCATA), for the evaluation of cosmetic creams. With TCATA, 22 semi-trained panelists were asked to check/uncheck all terms from a list of attributes (sticky, difficult to spread, easy to spread, white residue, fresh, smooth, waxy, greasy and oily) that applied/no longer applied to describe the sensation while applying the sample at each moment within a period of 60s. Differences in temporal profiles of the cosmetic cream samples were identified by a sign test and correspondence analysis on the citation frequencies for each attribute (Boinbaser et al., 2015). The results demonstrated the potential of TCATA for characterizing the temporal dynamics of other beauty care products such as the performance of perfumes or foaming ability of shampoo or face cleanser (Boinbaser et al., 2015).

Use of Text Mining Techniques to Conduct Sensory and Consumer Research

Conventional methods for data collection in consumer research can be time-consuming, labor-intensive, and costly. Meanwhile, most of these practices in sensory and consumer research place respondents in artificial settings that require them to think about their behaviors and adopt an analytical mindset (Köster, 2009). The growth of social media and e-commerce has opened the door to huge amounts of information freely generated without the intervention of researchers. These reviews, as the “voice of consumers”, can provide tremendous insights into consumer preferences from a larger consumer sample compared to conventional studies. The rapid development data mining and natural language processing (NLP) techniques has enable researchers to analyze massive amounts of unstructured text data collected from online sources such as webpages and bulletin boards. Text mining is an exploratory process that transforms unstructured text data to structured format so that patterns or insights can be explored. This process involves the notions of statistics, linguistics, and machine learning. The typical tasks include concept extraction, document summarization, text categorization and clustering, sentiment analysis, predicting models, etc.

Application in Sensory and Consumer Research

The application of online text mining techniques to analyze online review data shows its potential in facilitating sensory and consumer research on personal care products. For instance, Kim & Kang (2018) collected product reviews for 42 Korean beauty balm BB creams and 194 competing BB creams from other countries from MakeupAlley, a cosmetic review site. By performing ratio analysis of words, Latent Semantic Analysis (LSA), Labeled Latent Dirichlet Allocation (L-LDA), and polarity analysis on the review data, the authors extracted 40 attributes that differentiate the Korean BB creams from BB creams of foreign countries (Kim & Kang, 2018). For example, the results showed that Korean products have been perceived as having different colors, functions, and containers from the competitive products (Kim & Kang, 2018). This study represents a useful procedure for collecting and analyzing large-scale text data to understand consumer insights, while certain limitations have been acknowledged in the paper. For example, the polarity of the discriminating attributes needs further verification, and the data-processing procedure might not be suitable for other categories of products (Kim & Kang, 2018).

Another recent study conducted by Hamilton & Lahne (2020) demonstrated how methods from Natural Language Processing of review data can be applied to generate sensory terminology for whiskey. In this study, a total of 6,598 online reviews of international whiskies were collected from two prominent websites (Hamilton & Lahne, 2020). Using natural language processing techniques, review data was tokenized into individual words which were later filtered by only keeping potentially descriptive nouns, adjectives, and verbs (Hamilton & Lahne, 2020). Having removed terms with low frequency of presence in the reviews, the list was then grouped into a flavor wheel which consisted of 16 categories of descriptors using Correspondence Analysis and Agglomerative Hierarchical Clustering (Hamilton & Lahne, 2020). The authors compared the automated flavor wheel with an existing Scotch wheel, finding that they shared several categories of descriptors (Hamilton & Lahne, 2020). However, most of the terms generated using this method were 'non-negative' sensory descriptors and lacked off-flavor descriptors, which the author mentioned might not be sufficient for use in quality control (Hamilton & Lahne, 2020). Future research could focus on the analysis of reviews with low ratings and customer complaint data. Sensory research of personal care products such as cosmetics could benefit from this automated terminology development procedure, especially with a huge amount of product reviews available on different cosmetics retailing platforms. Even though online review research has the advantages of collecting data fast and generating insights from large-scale data, there are several limitations such as no experimental control, lacking information about demographics and challenges in processing unstructured text data, etc. An overview of the strengths and limitations of social media research can be found in (Vidal et al., 2018).

Text Mining Techniques

Text preprocessing is a must for text mining, which normally uses NLP techniques to perform tokenization, removal of stop words, lemmatization and stemming (Silge & Robinson, 2017). First, tokenization is the process that text is split into individual meaningful unite, called 'token' (Silge & Robinson, 2017). The token can be most often a word or a bigram (two words), or a trigram (three words) (Silge & Robinson, 2017). Next, in lemmatization or stemming, the prefixes and suffixes of words are removed, and words are reduced to their derive root word forma and meaning (Manning et al., 2009). Further, stop words that don't contribute to the overall meaning if the text such as “the”, “of”, “to” can be removed (Silge & Robinson, 2017). Part-of-Speech (POS) tagging is another NLP technique that is frequently used in text analysis. This technique assigned a tag to each token according to its part of speech, for example, denoting nouns, verbs, adjectives and so on, which help understanding the semantic meaning of the unstructured text (Kumawat & Jain, 2015). The preprocessing step also transforms the unstructured data to structured data for further statistical analysis. For instance, the tidytext package in R prepares unstructured text data in a tidy format, with each variable in a column, each observation in a row and each type of observation unit in a table (Wickham, 2014).

In the case of identifying key features of a product from online review data, word frequency, word association and clustering, topic modeling and sentimental analysis can be performed to uncover consumer insights for a category of products. Calculating word frequencies is a common task in text mining and it helps with extracting key words from the document. However, only examining how many times a word occur in a document might not be indicative enough about the importance of the word to a document, especially when comparing several documents, because some words might appear in all documents of comparison. In this case, to examine how important a word is to a document in a collection of documents, term frequency is normally adjusted for how rarely it is used. This parameter is called term frequency-inverse document frequency (tf-idf) and is normally calculated by multiple the frequency of a term in a document (tf) with inverse document frequency (idf). Idf is denoted as the formula below (Silge & Robinson, 2017).

$$df\ (term) = \ln(\frac{n(documents)}{n(documents\ containing\ term)})$$

Topic modeling is a method for unsupervised classification technique that has been actively used to detect main themes or latent topics in the text data (Blei et al., 2003). Latent

Dirichlet allocation (LDA) is a particularly popular method for fitting a topic model (Blei et al., 2003). It treats each document as a mixture of topics, and each topic as a mixture of words (Blei et al., 2003). Sentiment analysis, also known as opinion mining, is the process of analyzing text data to determine the emotional tone (positive, negative, or neutral) it carries. This technique has been used widely in a lot of fields where the importance of online reviews is emphasized such as consumer goods, politics, service industry such as movies, hotels, and restaurants (Mostafa, 2013; Nakayama & Wan, 2019; Park et al., 2020).

Research Objectives

Sensory and consumer research is essential for the development and innovation of beauty care products. Studying consumer emotions related to beauty care could provide valuable insights into a product beyond conventional liking measures. In addition, online product reviews provide large amount of information that represents "the voice of consumers". Text mining techniques could be used on this information to uncover consumer product experience.

This dissertation aims to explore consumer experience of beauty care products from two perspectives: liking and beyond liking (emotions) based on conventional sensory and consumer data and online product reviews. The objective of Chapter 2 was to develop an emotion lexicon that could be used to profile consumers' emotional responses to beauty care products in sensory and consumer tests. The objective of Chapter 3 was to identify the main sensory drivers of liking and emotions for hand creams in different consumer segments. Chapter 4 explored consumer experience for hand cream products from the "voice of consumers"-online product reviews, aiming to identify the key features that drive consumers' overall rating for hand creams.

References

ASTM Standard E1207 − 14. (2014). Guide for Sensory Evaluation of Axillary Deodorancy. ASTM International.

ASTM Standard E1490-19. (2019). Guide for Two Sensory Descriptive Analysis Approaches for Skin Creams and Lotions. ASTM International.

ASTM Standard E1958 − 20. (2020). Guide for Sensory Claim Substantiation. ASTM International.

ASTM Standard E2049 − 20. (2020). Guide for Quantitative Attribute Evaluation of Fragrance/Odors for Shampoos and Hair Conditioners by Trained Assessors. ASTM International.

ASTM Standard E2082 − 12. (2020). Guide for Descriptive Analysis of Shampoo Performance. ASTM International.

Aust, L. B., & Oddo, L. P. (1989). APPLICATIONS OF SENSORY SCIENCE WITHIN THE PERSONAL CARE BUSINESS: PART 1. Journal of Sensory Studies, 3(3), 181–185.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. Annu. Rev. Psychol., 58, 373-403.

Bestelmeyer, P. E. G., Kotz, S. A., & Belin, P. (2017). Effects of emotional valence and arousal on the voice perception network. Social Cognitive and Affective Neuroscience, 12(8), 1351–1358.

Bhumiratana, N., Adhikari, K., & Chambers, E. (2014). The development of an emotion lexicon for the coffee drinking experience. Food Research International, 61, 83–92.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. the Journal of machine Learning research, 3, 993-1022

Bloch, L. D., Valente, N. Y., Escudeiro, C. C., Sarruf, F. D., & Velasco, M. V. R. (2021). Chemical and physical damage affect the perceptions of hair attributes: A quantitative sensory assessment by a trained panel. Journal of Sensory Studies, 36(1), e12621.

Boinbaser, L., Parente, M. E., Castura, J. C., & Ares, G. (2015). Dynamic sensory characterization of cosmetic creams during application using Temporal Check-All-That-Apply (TCATA) questions. Food Quality and Preference, 45, 33–40.

Bourguet, C., Pêcher, C., Bardel, M.-H., Navarro, S., & Mougin, D. (2016). The use of an ethological approach to evaluate consumers' appreciation of luxury facial skincare and discriminate between products: A preliminary study. Food Quality and Preference, 50, 7–14.

Chrea, C., Grandjean, D., Delplanque, S., Cayeux, I., Le Calve, B., Aymard, L., Velazco, M. I., Sander, D., & Scherer, K. R. (2008). Mapping the Semantic Space for the Subjective Experience of Emotional Responses to Odors. Chemical Senses, 34(1), 49–62.

Churchill, A., & Behan, J. (2010). Comparison of methods used to study consumer emotions associated with fragrance. Food Quality and Preference, 21(8), 1108–1113.

Churchill, A., & Greenaway, R. (2018). Application of Descriptive Analysis to Non-Food Products. In S. E. Kemp, J. Hort, & T. Hollowood (Eds.), Descriptive Analysis in Sensory Evaluation (pp. 647–677). John Wiley & Sons, Ltd.

Coppin, G., & Sander, D. (2016). Theoretical Approaches to Emotion and Its Measurement. In Emotion Measurement (pp. 3–30). Elsevier.

David, O. A., David, D. O., Mogoase, C., Popescu, L. C., Giosan, C., & Pellegrino, A. (2019). Psychological effects and brain correlates of a rose-based scented cosmetic cream. Journal of Sensory Studies, 34(6), e12536.

Delarue, J., Masson, M., & Blumenthal, D. (2018). Consumer Evaluation of Nonfood Products. In Methods in Consumer Research, Volume 2 (pp. 411–426). Elsevier.

Desmet, P. M. A., & Schifferstein, H. N. J. (2008). Sources of positive and negative emotions in food experience. Appetite, 50(2–3), 290–301.

Dooley, L. M., Adhikari, K., & Chambers IV, E. (2009). A GENERAL LEXICON FOR SENSORY ANALYSIS OF TEXTURE AND APPEARANCE OF LIP PRODUCTS. Journal of Sensory Studies, 24(4), 581–600.

Dreyfuss, L. (2018). Cosmetic Products. In Methods in Consumer Research, Volume 2 (pp. 397–410). Elsevier.

Ekman, P. (1992). Are There Basic Emotions? Psychological Review, 4.

Ekman, P. (1999 Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. Science, 164(3875), 86-88.). Basic emotions. Handbook of cognition and emotion, 98(45-60), 16.

Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. Psychological science, 18(12), 1050-1057.

Ferdenzi, C., Delplanque, S., Barbosa, P., Court, K., Guinard, J.-X., Guo, T., Craig Roberts, S., Schirmer, A., Porcherot, C., Cayeux, I., Sander, D., & Grandjean, D. (2013). Affective semantic space of scents. Towards a universal scale to measure self-reported odor-related feelings. Food Quality and Preference, 30(2), 128–138.

Ferrarini, R., Carbognin, C., Casarotti, E. M., Nicolis, E., Nencini, A., & Meneghini, A. M. (2010). The emotional response to wine consumption. Food Quality and Preference, 21(7), 720–725.

Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. Journal of personality and social psychology, 57(2), 212.

Guest, S., Dessirier, J. M., Mehrabyan, A., McGlone, F., Essick, G., Gescheider, G., Fontana, A., Xiong, R., Ackerley, R., & Blot, K. (2011). The development and validation of sensory and emotional scales of touch perception. Attention, Perception, & Psychophysics, 73(2), 531–550.

Hamilton, L. M., & Lahne, J. (2020). Fast and automated sensory analysis: Using natural language processing for descriptive lexicon development. Food Quality and Preference, 83, 103926.

Hightower, C. A., & CHAMBERS IV, E. D. G. A. R. (2009). Descriptive sensory analysis of toothpaste flavor and texture using two sampling methods: brushing versus spoon tasting. Journal of sensory studies, 24(3), 301-316.

Izard, C. E. (2007). Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm. Perspectives on Psychological Science, 2(3), 260–280.

Izard, C. (1979). Facial expression, emotion, and motivation (pp. 31-49). Academic Press, New York, NY, USA.

Jiang, Y., King, J. M., & Prinyawiwatkul, W. (2014). A review of measurement and relationships between food, eating behavior and emotion. Trends in Food Science & Technology, 36(1), 15–28.

Kahraman, A., & Kazançoğlu, İ. (2019). Understanding consumers' purchase intentions toward natural-claimed products: A qualitative research in personal care products. Business Strategy and the Environment, 28(6),

Kao Corporation | Research on Skin Care with a Focus on Stimulation that Creates Pleasant Feelings. (2018a). Retrieved January 5, 2021, from https://www.kao.com/global/en/news/rd/2018/20181115-001/

Kao Corporation | Kao Confirms that Different Emotions are Evoked by Skin Care Products that Produce Different Textures During Application. (2018b) Retrieved January 5, 2021, from https://www.kao.com/global/en/news/rd/2018/20181129-001/

Kim, J. Y., Sung, Y. E., & Kim, K. O. (2013). Descriptive analysis of toothpaste flavor using novel sample preparation procedure. Journal of Sensory Studies, 28(2), 154-160.

Kim, S. G., & Kang, J. (2018). Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews. Information Processing & Management, 54(6), 938–957.

King, S. C., & Meiselman, H. L. (2010). Development of a method to measure consumer emotions associated with foods. Food Quality and Preference, 21(2), 168–177.

King, S. C., Meiselman, H. L., & Thomas Carr, B. (2013). Measuring emotions associated with foods: Important elements of questionnaire and test design. Food Quality and Preference, 28(1), 8–16.

Köster, E. P. (2009). Diversity in the determinants of food choice: A psychological perspective. Food Quality and Preference, 20(2), 70–82.

Kuesten, C., Chopra, P., Bi, J., & Meiselman, H. L. (2014). A global study using PANAS (PA and NA) scales to measure consumer emotions associated with aromas of phytonutrient supplements. Food Quality and Preference, 33, 86–97.

Kumawat, D., & Jain, V. (2015). POS Tagging Approaches: A Comparison. International Journal of Computer Applications, 118(6), 32–38.

Lagast, S., Gellynck, X., Schouteten, J. J., De Herdt, V., & De Steur, H. (2017). Consumers' emotions elicited by food: A systematic review of explicit and implicit methods. Trends in Food Science & Technology, 69, 172–189.

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.

Lazarus, R. S., & Lazarus, R. S. (1991). Emotion and adaptation. Oxford University Press on Demand.

Lee, I.-S., Yang, H.-M., Kim, J.-W., Maeng, Y.-J., Lee, C.-W., Kang, Y.-S., Rang, M.-J., & Kim, H.-Y. (2005). TERMINOLOGY DEVELOPMENT AND PANEL TRAINING FOR SENSORY EVALUATION OF SKIN CARE PRODUCTS INCLUDING AQUA CREAM. Journal of Sensory Studies, 20(5), 421–433.

Manning, C., Raghavan, P., & Schuetze, H. (2009). Introduction to Information Retrieval. 581.

Martins, V. B., Bordim, J., Bom, G. A. P., Carvalho, J. G. D. S., Parabocz, C. R. B., & Mitterer Daltoé, M. L. (2020). Consumer profiling techniques for cosmetic formulation definition. Journal of Sensory Studies, 35(2).

Meilgaard, M. C., Civille, G. V., & Carr, B. T. (2016). Sensory evaluation techniques (Fifth edition). CRC Press, Taylor & Francis Group, CRC Press is an imprint of the Taylor & Francis Group, an Informa business.

Meiselman, H. L. (2016). Emotion Measurement. In Emotion Measurement (pp. 645–697). Elsevier.

Moussour, M., Lavarde, M., Pensé-Lhéritier, A. M., & Bouton, F. (2017). Sensory analysis of cosmetic powders: personal care ingredients and emulsions. International journal of cosmetic science, 39(1), 83-89.

Mostafa, M. M. (2013). More than words: Social networks' text mining for consumer brand sentiments. Expert Systems with Applications, 40(10), 4241–4251.

Nakayama, M., & Wan, Y. (2019). The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews. Information & Management, 56(2), 271–279.

Painchault, T., Perrin, L., Loijens, L. W. S., Linden, X. van der, Theuws, H. M. M., & Zimmerman, P. H. (2020). Application of scientific measures to demonstrate the relaxing properties of Peony fragrance in hair care products. Journal of Sensory Studies, 35(2), e12560.

Parente, M. E., Ares, G., & Manzoni, A. V. (2010). APPLICATION OF TWO CONSUMER PROFILING TECHNIQUES TO COSMETIC EMULSIONS: CONSUMER PROFILING OF COSMETIC EMULSIONS. Journal of Sensory Studies, 25(5), 685–705.

Parente, M. E., Gámbaro, A., & Ares, G. (2008). Sensory Characterization of Emollients. Journal of Sensory Studies, 23(2), 149–161.

Parente, M. E., Manzoni, A. V., & Ares, G. (2011). EXTERNAL PREFERENCE MAPPING OF COMMERCIAL ANTIAGING CREAMS BASED ON CONSUMERS' RESPONSES TO A CHECK-ALL-THAT-APPLY QUESTION: PREFERENCE MAPPING OF ANTIAGING CREAMS BASED ON A CATA QUESTION. Journal of Sensory Studies, 26(2), 158–166.

Parente, M. E., Ochoa Andrade, A., Ares, G., Russo, F., & Jiménez-Kairuz, Á. (2015). Bioadhesive hydrogels for cosmetic applications. International Journal of Cosmetic Science, 37(5), 511–518.

Park, E., Kang, J., Choi, D., & Han, J. (2020). Understanding customers' hotel revisiting behaviour: A sentiment analysis of online feedback reviews. Current Issues in Tourism, 23(5), 605–611.

Plutchik, R. (1980). A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION. In Theories of Emotion (pp. 3–33). Elsevier.

Porcherot, C., Delplanque, S., Raviot-Derrien, S., Calvé, B. L., Chrea, C., Gaudreau, N., & Cayeux, I. (2010). How do you feel when you smell this? Optimization of a verbal measurement of odor-elicited emotions. Food Quality and Preference, 21(8), 938–947.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and Psychopathology, 17(03).

Richins, M. L. (1997). Measuring Emotions in the Consumption Experience. Journal of Consumer Research, 24(2), 127–146.

Russell, J. A., & Barrett, L. F. (1999). Core Affect, Prototypical Emotional Episodes, and Other Things Called Emotion: Dissecting the Elephant. 15.

Russell, J. A., & Pratt, G. (1980). A Description of the Affective Quality Attributed to Environments. 12.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A Single-Item Scale of Pleasure and Arousal. 10.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. Psychological bulletin, 115(1), 102.

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.

Schifferstein, H. N. J., & Desmet, P. M. A. (2010). Hedonic asymmetry in emotional responses to consumer products. Food Quality and Preference, 21(8), 1100–1104.

Schwartz, N. O. (1975). ADAPTATION OF THE SENSORY TEXTURE PROFILE METHOD TO SKIN CARE PRODUCTS. Journal of Texture Studies, 6(1), 33–42.

Spinelli, S., Masi, C., Dinnella, C., Zoboli, G. P., & Monteleone, E. (2014). How does it make you feel? A new approach to measuring emotions in food product experience. Food Quality and Preference, 37, 109–122.

Smith, C. A., & Lazarus, R. S. (1993). Appraisal components, core relational themes, and the emotions. Cognition & emotion, 7(3-4), 233-269.

Silge, J., & Robinson, D. (2017). Text mining with R: A tidy approach. " O'Reilly Media, Inc.".

Storm, C., & Storm, T. (1987). A taxonomic study of the vocabulary of emotions. Journal of personality and social psychology, 53(4), 805.

Sun, C., Koppel, K., & Chambers, E. (2014). An Initial lexicon of sensory properties for nail polish. International Journal of Cosmetic Science, 36(3), 262–272.

Talavera, M., & Sasse, A. M. (2019). Gathering consumer terminology using focus groups—An example with beauty care. Journal of Sensory Studies, 34(6).

Thomson, D. M. H., & Crocker, C. (2013). A data-driven classification of feelings. Food Quality and Preference, 27(2), 137–152.

Thomson, D. M. H., Crocker, C., & Marketo, C. G. (2010). Linking sensory characteristics to emotions: An example using dark chocolate. Food Quality and Preference, 21(8), 1117–1125.

Tomkins, S. (1963). Affect imagery consciousness: Volume II: The negative affects. Springer publishing company.

Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. Food Research International, 48(2), 893–908.

Vidal, L., Ares, G., & Jaeger, S. R. (2018). Chapter 6—Application of Social Media for Consumer Research. In G. Ares & P. Varela (Eds.), Methods in Consumer Research, Volume 1 (pp. 125–155). Woodhead Publishing.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. Psychological bulletin, 98(2), 219.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of personality and social psychology, 54(6), 1063.

Wilson-Mendenhall, C. D., Barrett, L. F., & Barsalou, L. W. (2013). Neural Evidence That Human Emotions Share Core Affective Properties. Psychological Science, 24(6), 947–956.

Wickham, H. (2014). Tidy data. Journal of statistical software, 59(1), 1-23.

Wortel, V. A. L., & Wiechers, J. W. (2000). Skin sensory performance of individual personal care ingredients and marketed personal care products. Food Quality and Preference, 11(1–2), 121–127.

Xing, H., Vaught, C., & Chambers, E. (2020). Using multiple data analysis methods to guide makeup remover wipe optimization in a design of experiments consumer home use test. Journal of Sensory Studies, 35(1).

Zevon, M. A., & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. Journal of personality and social psychology, 43(1), 111.

Chapter 2 - Development of An Emotion Lexicon for Beauty Care Products

Abstract

Emotions play a key role in the design and optimization of beauty care products. However, limited work has been published regarding the explicit measurement of consumer emotions associated with this category. The objective of this chapter was to develop an emotion lexicon that could be used to profile consumers' emotional responses to beauty care products in sensory and consumer tests. The lexicon was developed in four main steps, starting by sourcing terms from the "voice of consumers"- online product reviews. Then, a list of candidate terms that were either emotions or had emotional connotations extracted from the online reviews were categorized into different valences and refined by consumers who were beauty product users in interviews and surveys. Finally, the emotion lexicon was validated in an online test using product concepts of skincare, hair care and makeup as stimuli. The developed lexicon showed discriminating abilities between different categories of beauty care. The final emotion lexicon consists of 37 positive emotions and 20 negative emotions. The emotional space of beauty care was uncovered with Multidimensional Scaling (MDS) and Agglomerative Hierarchical Clustering (AHC). In addition, skincare and hair care were found to be more likely to evoke emotions related to self-wellbeing and sensory pleasure, while makeup products elicited social-oriented emotions and emotions related to vanity. A description of which emotion was considered more applicable to which category of skincare, hair care and makeup was presented. The validated emotion lexicon from this study is readily applicable to other emotion research for skincare, hair care and makeup. Text analysis of online reviews provides an effective way to gather consumer terminology for sensory and consumer research of beauty care.

Introduction

The emotional aspects of beauty products play a central role in determining consumer satisfaction and have been incorporated as the major factor into product design in the last few decades. This is because, on one hand, consumers who use beauty products nowadays are not only driven by the functions of the products, but more importantly, the emotional motivations such as the desire of feeling good, boosting self-esteem and social value, and improving physical attractiveness from the experience of beauty products. Beauty products evoke positive emotions to their users by conveying the perception of wellbeing and caring for oneself and reducing dissatisfaction with one's image (Apaolaza-Ibáñez et al., 2011). As a result, to stay competitive in the market, cosmetic companies need to keep investigating what emotions consumers expect, and tapping new emotions into the sensory profiles, claims, packages, positioning of their products. On the other hand, emotions can affect beauty. For example, a study conducted by a Japanese cosmetic company suggested that positive emotions evoked using skincare products resulted in better skin appearance in terms of radiance and clarity (Kao Corporation, 2018). The improved appearance was explained by the increased level of oxytocin within the body because of pleasant feelings created by tactile stimulation during skin care (Kao Corporation, 2018).

Understanding the emotional needs of consumers and their emotional associations with different formulations, claims and benefits offers innovation opportunities for cosmetic companies. Consumers' emotional responses to fragrances in different beauty care formulations have been measured in several studies using implicit methods including facial expression, physiological parameters such as heart rate, blood pressure, and electrodermal activity, as well as brain measures such as MRI (David et al., 2019; Painchault et al., 2020). Implicit measurements are normally used to demonstrate the emotional effect of an ingredient such as the soothing or relaxing effect of a fragrance. These methods have the advantages of measuring emotional responses without the conscious control of consumers, however, it has been found these measurement methods were unable to differentiate large numbers of positive emotions (Meiselman, 2016). In addition, implicit measurement or implicit combined with explicit measurement require special set-up instrumental wise, which makes them unable to test a large sample size as conventional consumer tests in the self-reported format (Lagast et al., 2017). As a result, these measures seem to be more suitable to be used in the initial stages of product development when selecting ingredients or screening prototypes with a small size of panelists.

Compared to implicit measurement, self-reported questionnaires using a list of emotional terms that can be rated or checked are so far the most used method to quickly measure product-related emotions, even though data collected are consciously emotional responses to products. To measure the affective feelings elicited by odors and fragrances, Chrea et al. (2008) developed a verbal questionnaire- Geneva Odor Scales (GEOS) with 36 emotion terms in 6 dimensions namely, pleasant feeling, unpleasant feeling, sensuality, relaxation, refreshment, and sensory pleasure. The GEOS has been later reduced to a shorter format for commercial testing and has been extend to a series of culture-specific Emotion and Odor Scales (EOSs) to 5 other cultures (Brazil, China, Singapore, United Kingdom, and the United States) (Ferdenzi et al., 2013; Porcherot et al., 2010). In addition, Guest et al. (2011) mapped the touch-related perception in sensory and emotional spaces, naming the Tactile Perception Task (TPT). This tool consists of 26 sensory attributes classified into four factors and 14 emotional attributes representing comfort and arousal dimensions, which can be used to measure emotions evoked by personal care products with a variety of textures. All these emotion lexicons were created based on a single sensory modality, either aroma or texture. However, the emotional experiences related to beauty care products can be the results of multi-sensory experience including visual, olfactory, tactile experience, potential functions and efficacy of the products, memories, and even social implications of the product usage, and interactions of all these factors (Parente et al., 2011). There is a need for a lexicon that can be used to profile consumers' emotional responses related to the whole usage experience of a beauty product. This type of lexicon can be used for testing beauty concepts, formulations, packages, branded and unbranded beauty care products. In a focus group study with beauty care product consumers, Talavera & Sasse (2019) gathered a list of terms that can be used to describe consumers emotions associated with makeup, skincare, and haircare categories. However, this list of emotions needs to be further validated before it can be used in consumer tests.

Emotion lexicons have been developed for many food and beverage products. These include domain-specific lexicons such as the EsSense Profile and the GEOs, or product-specific lexicons such as emotion lexicons for coffee, beer, blackcurrant, chocolate, and hazelnut spread (Bhumiratana et al., 2014; Chaya et al., 2015; Chrea et al., 2008; King & Meiselman, 2010; Ng et al., 2013). Most of these lexicons were developed by firstly sourcing potential terms either from psychological or psychiatric scales, such as the standardized mood questionnaires, or gathering consumer terminology by conducting focus groups or interviews. Then the initial list of candidate terms will be reduced by conducting surveys with consumers to identify the most relevant and discriminating terms related to a domain or product type. Despite that these lexicons showed adequate discriminating abilities and validity, none of these developments approached emotions from the perspective of consumers' actual language use data, which is crucial to fully reflect the terms consumers actively use in a given context. To address this, Gmuer et al. (2015) adopted a systematic and linguistic-based approach for the development of a food-related emotion lexicon for German consumers by sourcing terms from the most-used dictionaries and corpus in Germany. An alternative source of consumers' active language use data could be online product reviews. The popularity of social media and online purchases has opened the door to huge amounts of information about consumer product experience without the intervention of researchers. More and more consumers share their product experience online by writing product reviews. These reviews, as the "voice of consumers", can provide tremendous insights into consumers' experiences from a larger sample size compared to conventional sensory and consumer studies. Coupled with text analysis and natural language processing techniques, terminology that are actively used by consumers can be extracted from online product review data.

This study aimed to develop an emotion lexicon for the beauty care category starting from sourcing emotion terminology from online product reviews. From an application perspective, we defined "emotions" in a broad sense as any terms that could possibly describe feelings or subjective affective experiences related to beauty care products. These might include emotions, moods, and evaluative terms. Four studies were conducted for the whole development process which was described in **Figure 2.1**.

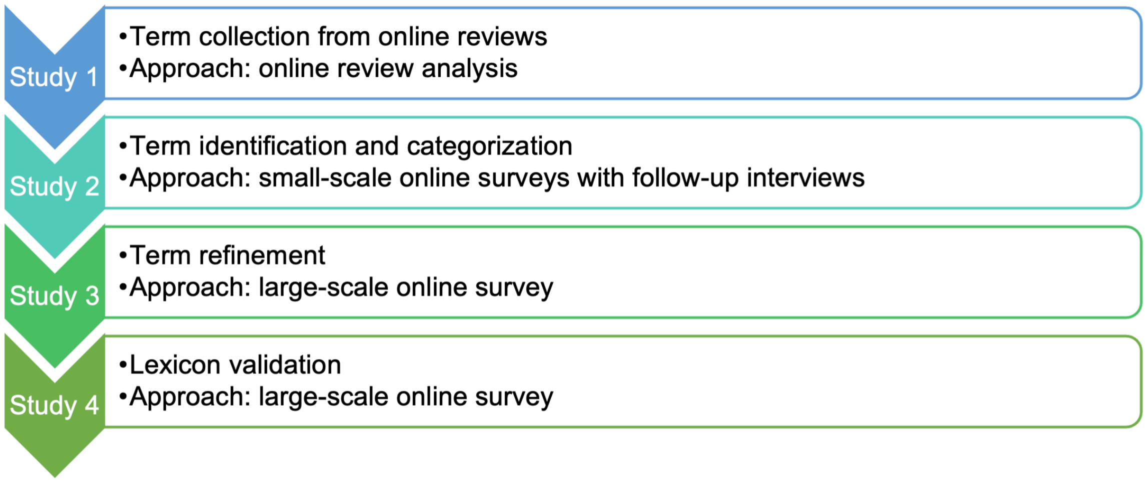


Figure 2.1. The development process for the beauty care emotion lexicon

STUDY 1: TERM COLLECTION

Study 1 aimed at sourcing the possible terms that could be used for constructing the emotion lexicon for beauty care. Emotions and terms with emotional connotations were identified via text analysis of online reviews collected from the website of Ulta Beauty (https://www.ulta.com/), which is the online store of a major beauty chain across the United States, and MakeupAlley (https://www.makeupalley.com), a beauty product review platform where consumers can write their reviews by giving ratings and free comments on a variety of makeup, haircare and skincare products.

Materials and Methods

Online Review Description

Online product reviews of skincare, haircare, and makeup were scraped from the two websites mentioned above using a web scraping software-ParseHub (Toronto, Ontario, Canada). Then, all review data collected were examined in Excel by automatically removing all duplications. This step resulted in a total of 88,123 online reviews for final analysis (Table 2.1). For each category of skincare, haircare and makeup, the reviews covered a wide range of overall ratings, product types, brands, prices, claims and sensory characteristics. The skincare category included reviews of facial treatment products (face lotion, cream, serum, gel, toner, exfoliator, anti-aging treatment, cleanser, and mask), eye products (eye cream and eye serum) and lip products (lip balm). The reviews for hair care were those of hair shampoos, conditioners, serums, sprays, and protective treatment products. The reviews for makeup products covered eye products (eye liner, mascara, eyebrow pencil, and eye shadow), face products (foundation, concealer, BB cream, contour, powder, primer), and lip products (lipstick, lip tint, lip plumper). Each review had the information of publication date, product comments and product rating, while only the product comments were used in this study.

Table 2.1. Description of the online review data for skincare, haircare, and makeup

	Skincare	Haircare	Makeup	Total
Number of products	35	36	28	99
Number of reviews	29,874	26,468	31,781	88,123

Data Analysis

Text analysis was carried out on the review data using the tidytext package (Silge & Robinson, 2016) in R (R Core Team, 2020). Within the tidy text framework, all comments were broken into individual words via a tokenization procedure, then, these data were transformed into a one-word-per-row format together with the frequency of each word. Stop words such as “the”, “of”, “to” and so forth in English were then removed from the dataset. To find the terms in our review data that were emotions or had emotional connotations, we compared all the words in our dataset with the words in sentiment lexicons (Table 2.2) provided within the tidytext package. The three sentiment lexicons we used were AFINN (Nielsen, 2011), Bing (Hu & Liu, 2004) and NRC (Mohammad & Turney, 2010, 2013). These sentiment lexicons are made up of English words that are either emotion words or words associated with emotions, and they are assigned scores for positive/negative sentiments or categorized into different basic emotion categories. Using the three sentiment lexicons as references, a word/term in our dataset would be kept if it appeared in any of these sentiment lexicons, otherwise, it would be removed from our dataset. This step helped us extract all terms/words that were associated with emotions or had sentimental meanings from the initial product reviews. Further steps were taken on the list of terms by firstly lemmatizing all words to their basic form, grouping the same words after lemmatization, and then converting all words to their corresponding adjectives or nouns using the Merriam-Webster Dictionary. The terms went through another round of manual examination conducted by two researchers: only terms that could possibly represent emotions, feelings, moods, or evaluative terms were kept. Finally, the term list was reduced by grouping synonyms using Merriam-Webster Thesaurus.

Table 2.2. Brief overview of the three sentiment lexicons: AFINN, Bing, and NRC

Sentiment Lexicon	Number of Terms	Term Categorization	Lexicon Access
AFINN (Nielsen, 2011)	2,477	Sentiment score from -5 to 5	http://www2.imm.dtu.dk/pubdb/pubs/6010-full.html
Bing (Hu & Liu, 2004)	6,786	Positive and negative	https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html
NRC (Mohammad & Turney, 2010, 2013)	14,182	Positive, negative, anger, anticipation, disgust, fear, joy, sadness, surprise, and trust	http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

Study 1 Results

Text analysis on review data and manual selection resulted in a primary collection of 330 terms including emotions, moods, evaluative terms, and other descriptive terms with emotional connotations. 233 out of the 330 candidate terms were reported in previous practical and theoretical emotion research (Chrea et al., 2008; King & Meiselman, 2010; Laros & Steenkamp, 2005; Talavera & Sasse, 2019). To make the starting list of terms as extensive as possible, twelve other terms (adult, approachable, awake, frizzy, frumpy, naked, old, put-together, self-conscious, unconfident, unkempt, unprofessional) generated by Talavera & Sasse (2019) from a beauty care focus group study but not mentioned in the online review data were added to our initial list. Overall, this study gathered a total of 342 candidate terms that could be used to describe emotions/feelings related to beauty care products. The list of candidate terms can be found in Appendix A.

STUDY 2: TERM IDENTIFICATION AND CATEGORIZATION

A total of 342 terms gathered in Study 1 was the starting point for the development of the emotion lexicon. Study 2 was designed to identify the relevant terms from the set of 342 terms by eliminating terms irrelevant to describe consumers' emotions related to beauty care products and terms that had ambiguous valences for consumers, as well as grouping synonyms in the context of beauty product experiences. This was achieved by conducting a small-scale online survey with follow-up interviews with beauty care consumers.

Materials and Methods

Participants Recruitment

Two groups of female participants (N1=22, N2=21) who were at least 18 years old (Table 2.3) were recruited from Kansas City area via the consumer database of Sensory and Consumer

Research Center at K-State Olathe. No significant differences between the two groups were found in terms of age (alpha=0.05). To qualify for the study, participants had to be medium-heavy users of all three categories of skincare, hair care and makeup. This meant they must typically use at least 3-4 different types of skincare products daily, wear makeup products at least 4 times a week, wear at least 3 different makeup products when doing makeup, and use at least 2 hair care products beside shampoos daily.

Table 2.3. Summary of participants' age for group 1 and group 2

Age	Group 1 (N=21)	Group 2 (N=22)
18-24	19%	23%
25-34	19%	23%
35-44	19%	23%
45-54	24%	18%
55+	19%	14%

Online Survey

The 342 terms collected from Study 1 were randomly divided into two lists of 171 terms. Each group of participants only responded to one of the two lists. In each questionnaire, participants were asked to think about the times when they had good or bad experiences using beauty care products (skincare, hair care or makeup), and then sort all terms from the list into 'positive emotions associated with beauty care products', 'negative emotions associated with beauty care products', and 'not relevant to beauty care products'. Data collection was completed using Compusense Cloud Software (Compusense, Inc., Guelph, Ontario, Canada).

Follow-up Interviews

Follow-up interviews were conducted with a selected group of participants to further understand consumers' rationales in term categorization, as well as the appropriateness to beauty care testing of the terms, especially for those terms that were categorized as negative or positive by low percentages of participants. This would help provide guidance for further keeping or eliminating certain terms and grouping synonyms.

Based on their responses to the online survey, twelve participants were selected for the follow-up interviews. These participants reported mixed and diverse opinions in categorizing the emotion terms. Each of the twelve participants took part in a 20-min one-on-one interview session conducted using Zoom (San Jose, CA) by a trained moderator. All participants consented and agreed to have the session recorded. The interviews began by showing the interviewees a list of terms together with their categorizations from the survey (the selection of terms showing was different for each participant depending on their responses from the online survey). Then, they were asked to explain the reasons why they categorized each of these terms in the way they did and to provide a scenario when using a specific beauty care product would make them feel that way. To end the session, each participant was asked about their opinions on a few synonym groups in the context of beauty care experience.

Data Analysis

All data analysis was conducted with Microsoft Excel (Redmond, WA). Terms sorted by more than 60% of participants as 'not relevant to beauty care products' were grouped as the irrelevant to beauty care group. For the rest of the terms, binomial test ($p=1/2$, one-tail, $\alpha=0.05$) was conducted on each term to compare the frequency of participants who categorized it as positive with frequency of being categorized as negative. Terms that showed significant differences were grouped into either positive group or negative group. Terms of no significant differences were considered as the ambiguous group. Terms in the irrelevant group and ambiguous group were further examined by participants in the follow-up interviews before being eliminated. For terms categorized into the positive or negative groups, potential synonyms within the group were identified by the researchers using a thesaurus (Merriam-Webster). Synonym terms were further evaluated during the interviews for further term reduction.

Study 2 Results

Seventy-eight terms were categorized as irrelevant by more than 60% of participants. Seventy-seven out of these terms were removed from the list. The only term that was kept was 'disgusted' as it seemed to be an important emotion related to negative experience of beauty care for certain consumers, even though it was reported to be irrelevant by a high percentage of participants. In the follow-up interviews, interviewees explained they would feel 'disgusted' when their makeup products did not work well, for example, when the mascara made their eyelashes so hard and brittle, or when any beauty care products had bad scents. For terms that were considered relevant by the participants, 153 terms were identified as positive terms and 91 were identified as negative terms according to the binomial tests. 20 terms were categorized as ambiguous terms in valence and were removed from the list. Some of these ambiguous terms were reported in previous research. For example, the term 'mature' was reported to be a positive emotion elicited by makeup/cosmetics (Talavera & Sasse, 2019). However, results from the online survey in this study suggested there were no significant differences between the number of participants who categorized this term as positive with those who thought it was negative. This suggested that interpretation of this term may be different among consumers. The follow-up interviews indicated that participants who perceived 'mature' as a positive emotion interpreted this term as sophisticated and professional, and they were typically young participants. On the other hand, those who thought 'mature' was a negative emotion as they assumed it meant old to them. As 'sophisticated' and 'old' were already in the final list of terms, 'mature' was eliminated. 'Flashy', 'glitter' and 'shiny' also meant controversially to different participants. Some participants in the interviews associated 'flashy', 'glitter' and 'shiny' with metallic colors that they tried to avoid. Some mentioned that when they felt shiny, their skin/hair were oily. Participants who considered 'flashy', 'glitter' and 'shiny' as positive terms associated these terms with trendy makeup products. According to the follow-up interviews, these three terms seem more used to describe the appearances of the cosmetic products rather than affective feelings. Another 14 terms such as 'soft', 'bold', 'crusty', 'flaky', 'frizzy', 'damaged', and 'cracked' that were more often used as descriptors for product features or sensory effects on skin/ hair were also eliminated. In the end, synonyms (from the positive and negative groups) identified using the thesaurus (Merriam-Webster) were examined by participants in the interviews for their semantic meanings in the context of beauty care usage experience. Terms reported to have similar meanings were grouped or eliminated. Altogether, these steps resulted in a total of 74 positive terms and 47 negative terms (Table 2.4).

Table 2.4. Positive and negative emotion terms identified from Study 2

Positive Terms (n=74)	Accomplished	Fascinated	Prestigious
	Active	Free	Pretty/Beautiful
	Adult	Fresh	Professional
	Adventurous	Fun	Proud
	Amazed	Genuine	Put-together/Polished
	Appreciative	Glamorous	Ready
	Approachable	Glowing	Reassured
	Attractive	Happy	Refreshed/Rejuvenated/Energized
	Awake	Healthy	Relaxed
	Awesome/Wonderful	Hopeful/Promising	Relieved
	Balanced	Impressed	Responsible
	Boosted	Inspired	Rewarded

	Calm	Loving	Romantic
	Carefree	Loyal	Secure
	Chic	Luxurious	Self-indulgent
	Clean	Motivated	Sexy
	Comfortable	Natural	Soothed
	Confident	Neat	Sophisticated
	Cool	Nourished	Special
	Creative	Outstanding	Strong
	Desire/Crave	Pampered	Superior
	Ecstatic	Passionate	Trust
	Effortless	Perfect/Flawless	Vibrant
	Elegant	Pleasant	Youthful
	Excited	Pleased/Satisfied	
Negative Terms (n=47)	Annoyed/Irritated	Frustrated	Strange
	Bad	Gross	Stressed
	Bland	Incomplete	Tired
	Bored	Indifferent	Ugly
	Cautious	Inferior	Uncomfortable
	Clueless	Insecure	Unconfident
	Concerned/Worried	Intimidated	Unhappy
	Confused	Lazy	Unhealthy
	Disappointed	Old	Unimpressed
	Disgusted	Overwhelmed	Unnatural
	Dissatisfied	Regret	Unpleasant
	Dull	Ridiculous	Unprofessional
	Duped/Deceived	Sad	Upset
	Embarrassed	Self-conscious	Vain
	Flawed	Skeptical	Vulnerable
	Frumpy	Sloppy/Messy/Unkempt	

STUDY 3- REFINEMENT OF THE TERM LIST

This study was designed to refine the list of emotion terms generated from Study 2 by identifying the most relevant terms to both positive and negative experiences for each of the three beauty care categories (skincare, hair care and makeup). To achieve this goal, the term list consisting of 74 positive terms and 47 negative terms were evaluated by users of skincare, hair care, and makeup separately in online surveys.

Materials and Methods

Participants Recruitment

Three groups of female participants (N=100 per group) who were at least 18 years old were recruited across the United States through an online market research agency (Dynata, Shelton, CT). No significant differences between three groups were found in terms of age (alpha=0.05). The three groups were regular users of skincare, hair care and makeup respectively. Participants in the skincare group had to use at least 2 types of skincare products in their daily regime. Hair care users were those who had a daily routine on hair care and used hair care products other than shampoos and conditioners. Participants of the makeup group were females who wore makeup at least two times per week and used at least 2 types of makeup products while wearing makeup.

Questionnaire

Separate questionnaires were developed for users of skincare, hair care and makeup. Each group only completed one questionnaire of the designated beauty care category. The emotion elicitation method was adapted from (King & Meiselman, 2010a). In the category-specific questionnaire, participants were first asked to describe two of their favorite and least favorite skincare/hair care/makeup products, and then respond to check-all-that-apply (CATA) questions to select all terms that described their emotions/feelings related to their favorite/least favorite products. If participants did not have any least favorite skincare/hair care/makeup products, they were then asked to describe any two products that had ever gave them negative experiences. The 74 positive terms were shown to participants in three pages, and the 47 negative terms were listed in two pages. The pages as well as the positions of terms on each page were randomized between participants. Data collection was completed by Dynata (Shelton, CT) using their online survey platform.

Data Analysis

Poor responses were first removed from the raw data. These included CATA data from participants who reported they did not have a favorite or least favorite products or wrote random information in the open-ended comments. The frequency count and percentage for each emotion being used by each group of consumers to describe their favorite and least favorite products were calculated. Positive terms that were checked by ≥ 20% of consumers in each category were kept and the cut-off point for negative terms was set to be 15% check rate. The cutoff lines (20% and 15%) were selected to cover the diversity and number of terms consumers used to describe their positive and negative experiences related to the three categories of beauty care. Twenty (20%) for positive terms and 15% for negatives were in-between the medians and 3rd quartiles of term selection rate respectively in our dataset. Twenty (20%) has been used as a cut-off point for identifying emotions related to food products in the development of the EsSense Profile (King & Meiselman, 2010). After eliminating the terms with low frequency of selection, chi-square test of independency was conducted to examine if each individual term could differentiate product categories. To better visualize the relationships between emotions and product categories, correspondence analysis (CA) was conducted on the total counts of positive terms and negative terms separately across skincare, hair care and makeup. Based on the CA maps, the list was further reduced by grouping highly correlated terms/terms of co-occurrence. Data analysis was conducted using XLSTAT (Addinsoft, New York, NY). A significance level of 0.05 was considered in this study.

Study 3 Results

Based on the cut-off criteria for term selection, 40 positive terms and 27 negative terms (Table 2.5 and Table 2.6) were retained from the initial list in the online survey. This meant that these positive terms were used by ≥ 20% participants in at least one of the three groups of consumers to describe their favorite skincare or hair care or makeup products. Meanwhile, ≥ 15% participants in at least one of the three questionnaires have used these 27 negative terms to describe their least favorite products. Chi-square test of independence performed on each of the 67 terms across skincare, hair care and makeup indicated 27 positive emotions and 7 negative emotions were able to discriminate between product categories. These terms were bolded terms with star labels in Table 2.5 and Table 2.6.

Table 2.5. Positive emotion terms selected for skincare, hair care and makeup

Positive Emotions	Skincare (N=98)	Hair Care (N=99)	Makeup (N=100)
Accomplished*	8%	21%	11%
Adult*	7%	10%	23%
Approachable*	8%	14%	24%
Attractive*	23%	33%	51%
Awake*	23%	7%	18%
Awesome/Wonderful	15%	20%	17%
Balanced	23%	18%	15%
Calm*	24%	11%	15%
Carefree	14%	23%	15%
Clean*	55%	41%	14%
Comfortable	34%	30%	38%
Confident*	24%	36%	48%

Fresh*	69%	38%	44%
Fun*	8%	13%	24%
Glamorous*	18%	21%	46%
Glowing*	38%	17%	37%
Happy	33%	35%	43%
Healthy*	30%	32%	16%
Luxurious	24%	27%	30%
Motivated*	9%	6%	24%
Natural*	48%	29%	32%
Neat	21%	17%	14%
Nourished*	44%	33%	6%
Pampered*	36%	28%	18%
Perfect/Flawless	22%	19%	32%
Pleasant*	35%	32%	16%
Pleased/Satisfied	55%	51%	39%
Pretty/Beautiful*	31%	34%	60%
Proud	11%	17%	20%
Put-together/Polished*	20%	31%	42%
Ready	29%	27%	30%
Refreshed/Rejuvenated/Energized*	34%	31%	13%
Relaxed*	27%	34%	15%
Relieved*	21%	14%	5%
Sexy*	10%	20%	31%
Soothed*	45%	15%	10%
Sophisticated*	9%	16%	22%

Special	8%	14%	20%
Vibrant	37%	23%	28%
Youthful*	24%	6%	26%

**Bolded terms were able to discriminate between product categories (Chi-square test of independence, alpha=0.05)*
Terms selected by ≥ 20% of participants in each category were highlighted and considered applicable to the corresponding category.

Table 2.6. Negative emotion terms selected for skincare, hair care and makeup

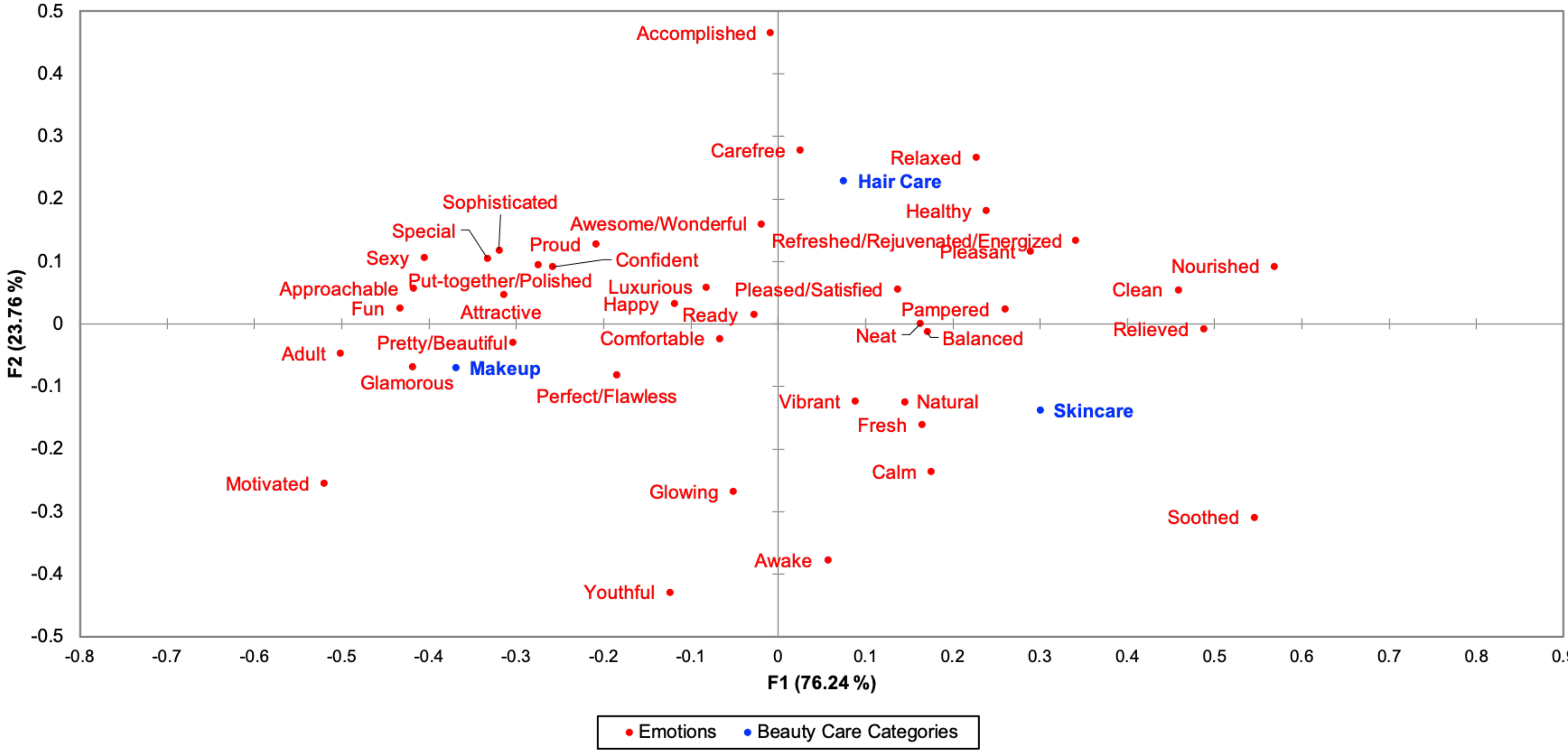
Negative Emotions	Skincare (N=85)	Hair Care (N=85)	Makeup (N=87)
Annoyed/Irritated	34%	26%	26%
Bad	15%	13%	18%
Bland	19%	14%	11%
Disappointed	42%	38%	32%
Dissatisfied	45%	49%	38%
Dull	22%	24%	11%
Duped/Deceived*	21%	13%	7%
Flawed*	16%	15%	31%
Frumpy	8%	14%	15%
Frustrated	29%	34%	31%
Gross	15%	18%	20%
Incomplete	19%	12%	11%
Insecure	14%	8%	21%
Old	12%	7%	16%
Regret*	32%	20%	14%

Ridiculous*	5%	2%	21%
Self-conscious*	19%	13%	29%
Sloppy/Messy/Unkempt*	16%	33%	28%
Ugly	11%	15%	23%
Uncomfortable	31%	26%	34%
Unconfident	18%	22%	29%
Unhappy	21%	25%	21%
Unhealthy	20%	22%	10%
Unimpressed	35%	29%	21%
Unnatural*	19%	14%	31%
Unpleasant	27%	24%	30%
Upset	25%	15%	13%

**Bolded terms were able to discriminate between product categories (Chi-square test of independence, alpha=0.05)*
Terms selected by ≥ 15% of participants in each category were highlighted and considered applicable to the corresponding category.

Correspondence Analysis (CA) biplots (Figure 2.2) showed how the three product categories were similar and different with each other in terms of positive and negative emotions they elicited. The first dimensions of both CA map (positive and negative) explained more than 70% of the total variation of the data, and they both separated makeup products with skincare and hair care products. Skincare and hair care were mainly differentiated by the second dimensions of correspondence analysis which only accounted for less than 25% of the total variation in cases of both positive and negative emotions. This indicated that skincare and hair care products were more similar in the emotions they elicited than makeup products did, especially for positive emotions. Compared with skincare and hair care, positive experiences of makeup products were more likely to elicit social-oriented emotions such as ‘approachable’, ‘adult’, ‘fun’, ‘motivated’, ‘sophisticated’, ‘sexy’, and ‘pretty/ beautiful’ and ‘attractive’. Skincare and hair care, on the other hand, seemed to elicit emotions related to self-wellbeing and sensory pleasure such as ‘clean’, ‘healthy’, ‘nourished’, ‘pampered’, ‘pleasant’, ‘refreshed/rejuvenated/energized’ and ‘relaxed’. ‘Accomplished’, ‘awesome/wonderful’, and ‘carefree’ were the terms more applicable for hair care. Terms more applicable for skincare were ‘awake’, ‘balanced’, ‘calm’, and ‘natural’, ‘relieved’ and ‘soothed’ which seemed to be related to the sensory and functional effects of the products. The differences in positive emotions between product categories could be due to the different functions the product category serves. Negative experiences of the three categories of beauty care tend to elicit similar negative emotions between product categories. There were only few terms specific to each category: feeling ‘duped/deceived’ and ‘regret’ were more related to describe bad experiences regarding skincare, ‘sloppy/messy/unkempt’ was more applicable for hair care, and ‘flawed’, ‘ridiculous’, ‘self-conscious’, and ‘unnatural’ were more specific for makeup products.

(a) Positive Emotions



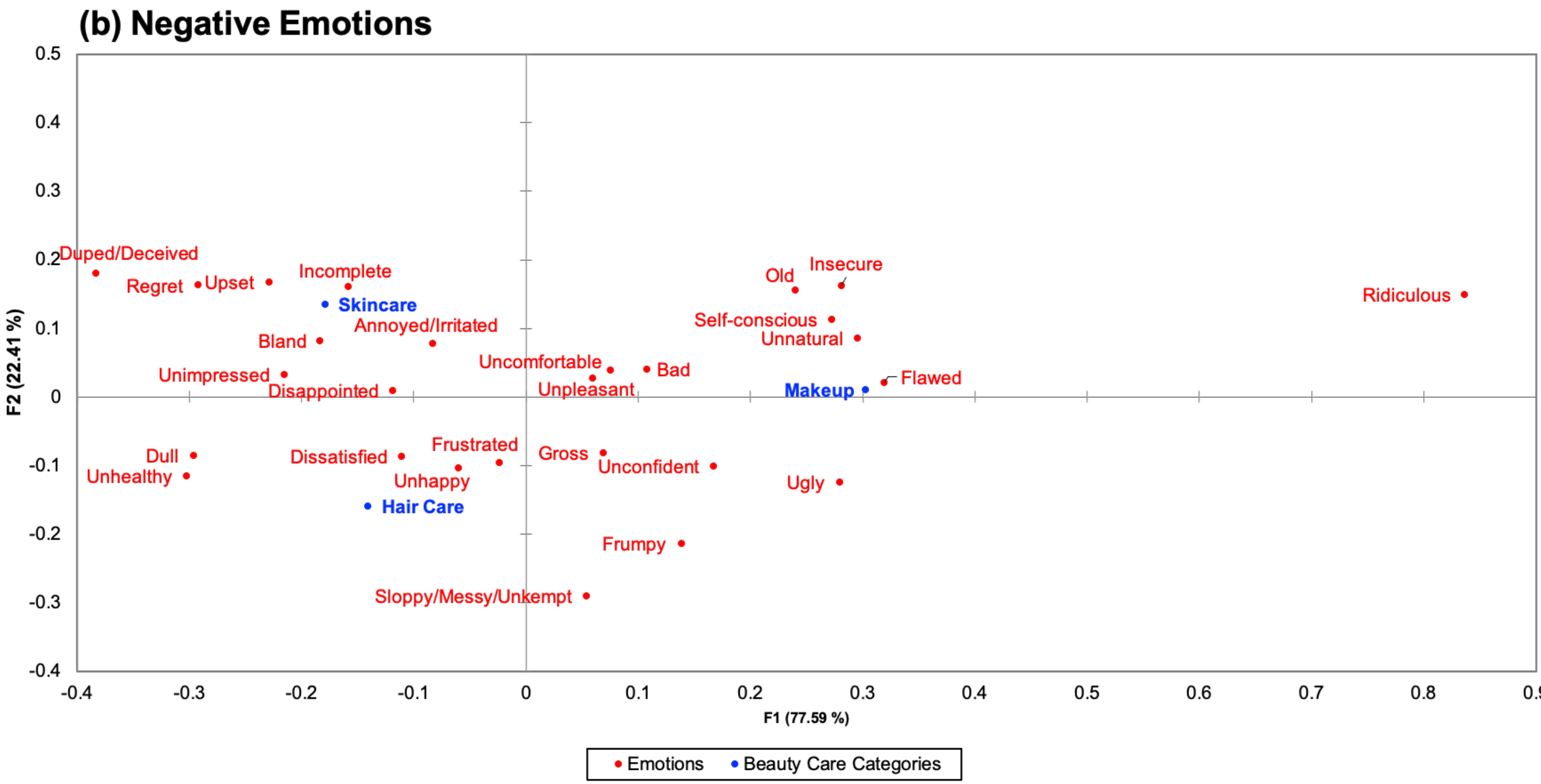


Figure 2.2. Correspondence analysis biplot from CATA total frequency counts of positive (a) and negative (b) emotions associated with favorite and least favorite skincare, hair care and makeup products

Based on the CA plots, terms that were close to each other on the map and had similar semantic meanings were grouped together to reduce to length of the lexicon. These groupings were 'pretty/beautiful' with 'attractive', 'calm' with 'soothed', 'bad' with 'unpleasant', 'insecure' with 'unconfident', 'self-conscious' with 'uncomfortable', 'bland' with 'dull'. This study completed the development of the initial emotion lexicon containing 38 positive terms and 23 negative terms for beauty care products (Table 2.7).

Table 2.7. The initial emotion lexicon for beauty care products

Positive Terms (n=38)	Accomplished	Glamorous	Pretty/Beautiful/Attractive
	Adult	Glowing	Proud
	Approachable	Happy	Put-together/Polished
	Awake	Healthy	Ready
	Awesome/Wonderful	Luxurious	Refreshed/Rejuvenated/Energized
	Balanced	Motivated	Relaxed
	Calm/Soothed	Natural	Relieved
	Carefree	Neat	Sexy
	Clean	Nourished	Sophisticated
	Comfortable	Pampered	Special
	Confident	Perfect/Flawless	Vibrant
	Fresh	Pleasant	Youthful
	Fun	Pleased/Satisfied	
Negative Terms (n=23)	Annoyed/Irritated	Old	
	Bad/Unpleasant	Regret	
	Bland/Dull	Ridiculous	
	Disappointed	Self-conscious/Uncomfortable	
	Dissatisfied	Sloppy/Messy/Unkempt	
	Duped/Deceived	Ugly	
	Flawed	Unhappy	
	Frumpy	Unhealthy	
	Frustrated	Unimpressed	
	Gross	Unnatural	
	Incomplete	Upset	
	Insecure/Unconfident		

STUDY 4- VALIDATION OF THE EMOTION LEXICON FOR BEAUTY CARE

Study 1, Study 2 and Study 3 developed an initial emotion lexicon consisting of 38 positive terms and 23 negative terms for beauty care products. Using beauty product concepts as testing stimuli in this study, an online survey was conducted to examine the discriminating ability of the initial emotion lexicon. The lexicon was further refined and recommendations on the application of this lexicon to each of the three beauty care categories was provided.

Materials and Methods

Participant Recruitment

One-hundred and fifty female participants who were at least 18 years old were recruited from Kansas City area via the consumer database of Sensory and Consumer Research Center at K-State Olathe. These participants were regular consumers of beauty care who used all three categories: skincare, hair care and makeup. They typically used at least two different skincare products daily; they used at least one hair care product other than shampoos and conditioners in their hair care regimen; and they used at least two different (makeup) products when applying makeup (at least 2 times per week).

Beauty Product Concepts

Concepts adapted from four commercial products were selected (Figure 2.3) as stimuli to elicit consumers' emotional responses. The four concepts were from generic or unfamiliar brands to elicit diverse emotional responses. These concepts included one skincare product, one hair care product, and two makeup products. They were Glow Recipe Avocado Melt Retinol Sleeping Mask (Glow Recipe, New York, NY), Klorane Dry Shampoo with Oat Milk for All Hair Types (Laboratories Pierre Fabre, Paris, France), 100% Pure Fruit Pigmented Ultra Lengthening

Mascara (Purity Cosmetics Inc., San Jose, CA), and Kaja Vacay Shine Lip Balm Oil (MBX Corporation, San Francisco, CA). Each product concept showed the name of the product (excluding brand names), its function, the key ingredients, and the usage instructions.

Online Survey

The online survey was conducted in Fall (October) 2020. A rate-all-that-apply (RATA) format was used for collecting emotional responses to different beauty product concepts in the online questionnaire (Ares et al., 2014). Participants evaluated the four product concepts sequentially with the order of concept presentation following a completely randomized design. In the online survey, participants were asked to first rate their overall liking for each product concept on a 9-pt hedonic scale and then rate-all-that-apply to a list of 61 emotion terms developed from study 3. The RATA procedure was implemented by having consumers check-all-all-apply (CATA) terms they considered applicable to describe their emotions first, and then rate the intensity of the terms that they checked on a 3-pt scale (1=low, 2=medium, 3-high) (Ares et al., 2014). For the CATA step, the 61 emotion terms were randomly distributed on three pages with 21 terms on the first page and 20 terms on the second and third page. Within each page, the positions of emotion terms were randomized between participants. Data collection was completed using Compusense Cloud Software (Compusense, Inc., Guelph, Ontario, Canada). It took participants about 15-20 minutes to complete the survey.

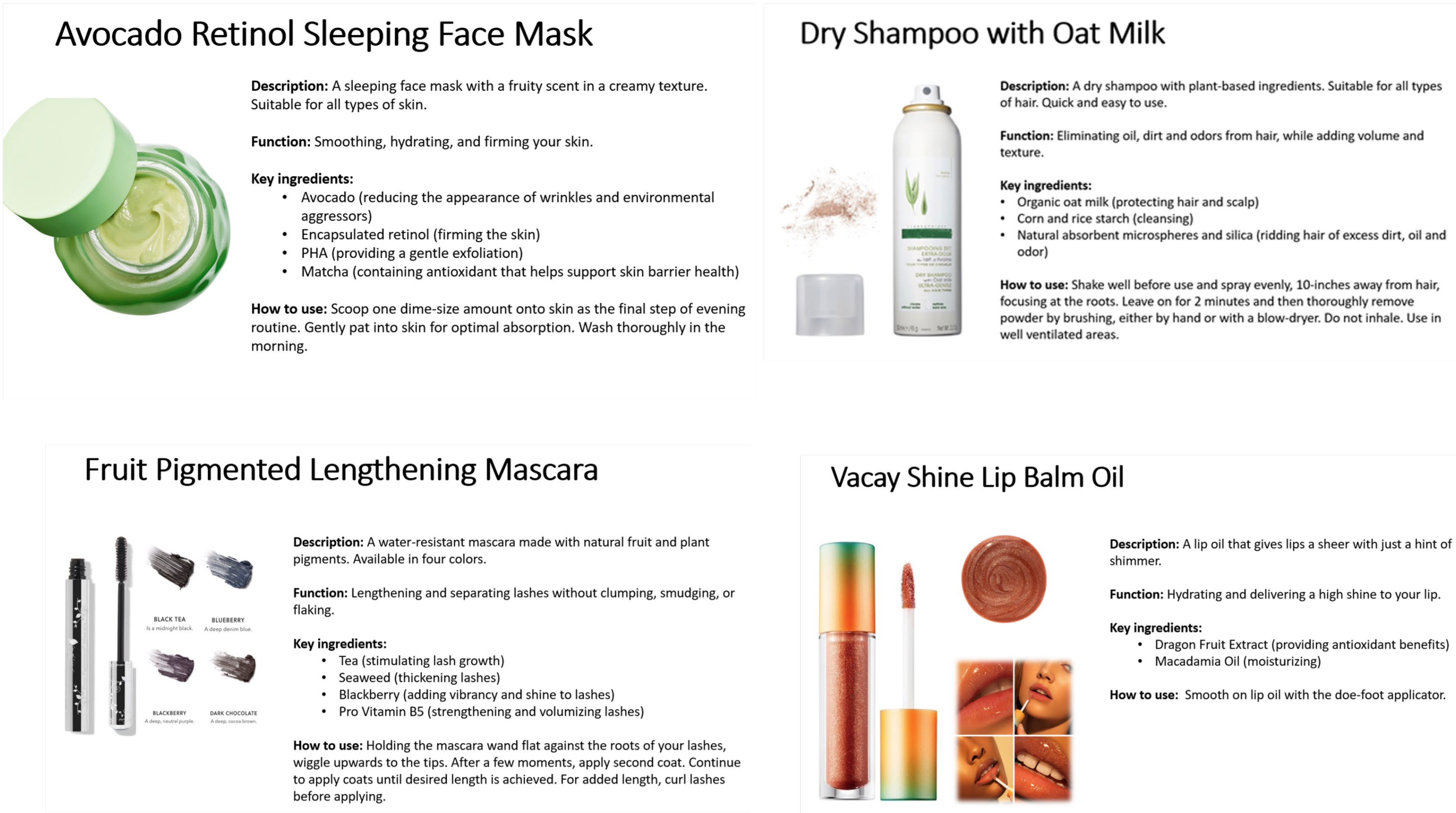


Figure 2.3. Product concepts used in the online survey

Data Analysis

Hedonic liking scores for the four concepts were compared by conducting analysis of variance (ANOVA) considering sample as fixed effect and participant as random effect. Tukey's HSD post hoc test was performed on any significant sample effect detected by ANOVA.

The discriminating ability of the initial emotion lexicon was examined by analyzing the RATA data in two ways: 1) considering RATA data as the binary CATA data, and frequency of selection for each term was used; 2) considering RATA data as continuous intensity scores by expanding the scale to 0-3 pt where 0 represented the term not being selected (Meyners et al., 2016). Meyners et al. (2016) suggested that despite the possibilities of violation of the required assumptions in parametric analysis (e.g. normal distribution and equidistant scale), results from parametric tests were similar to those from non-parametric (Friedman's) tests. Cochran's Q test was performed on the RATA frequency data to identify significant differences for each of the emotion terms among concepts. Sheskin's critical difference was used for multiple pairwise comparisons of any significant terms detected by Cochran's Q.

ANOVA was performed on the RATA intensity data for each emotion term considering sample as fixed effect and participant as random effect. Tukey's HSD post hoc test was performed on any significant terms detected by ANOVA. To visualize the relationship between product concepts and emotion terms, principal component analysis (PCA) was conducted on the mean RATA intensity values of the all the significant/discriminating emotion terms.

For term refinement purposes, percentage of selection of each emotion in each product concept was adjusted by only considering participants who liked the product concept (overall liking>5) for positive emotions, and those who were neutral or dislike the product concept (overall liking ≤5) for negative emotions. The final emotion lexicon was developed by further eliminating four terms with low adjusted usage rate in this study.

Pearson's correlations between all pairs of emotions were calculated using individual ratings. A two-dimensional representation of the emotional space of the developed lexicon was uncovered by multidimensional scaling (MDS) on similarity matrix (correlation matrix). Agglomerative Hierarchical Clustering (AHC) was conducted based on correlation distance (1-correlation matrix) using unweighted pair group average linkage. All data analysis was conducted using XLSTAT (Addinsoft, New York, NY). A significance level of 0.05 was considered in this study.

Study 4 Results

Discriminative Ability of the Emotion Lexicon

No significant differences were found for overall liking among these concepts (Figure 2.4) from ANOVA and all four product concepts were well liked by participants. This was not surprising as in concept testing, consumers normally rate the products as an expected product, and concept tests were found to be rated higher than actual product testing (Ares & Varela, 2018).

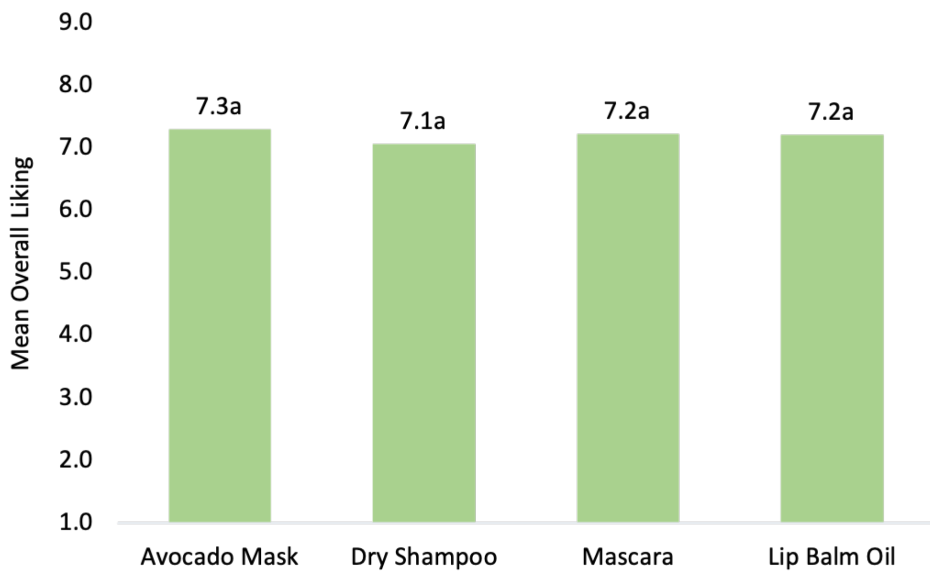


Figure 2.4. Overall liking scores for the four product concepts. Means with the same letter designation were not statistically different (alpha=0.05)

The emotion lexicon showed adequate discriminating abilities among product concepts of different beauty care categories no matter when treating RATA responses as intensity data or frequency data. The tested concepts received significantly different intensity ratings for 32 out of 38 positive terms and 6 out of 23 negative terms ($p<0.05$) (Table 2.8). When considering RATA frequencies, Cochran's Q tests identified a total of 37 out of 61 terms including 30 positive terms and 6 negative terms that significantly differentiated the four product concepts ($p< 0.05$) (Table 2.9). The only term that was found significant in RATA intensity analysis but non-significant in RATA frequency analysis was 'perfect/flawless.

PCA on the mean RATA intensities of the significant emotions (Figure 2.5) was conducted to visualize how these four concepts were differentiated overall in the emotions they elicited. The first dimension of the PCA explained 52.46% of the total variation, which mainly separated the makeup concepts (lip balm oil and mascara) with the group of skincare (avocado mask) and hair care (dry shampoo). The second dimension explaining 38.10% of the variation differentiated the concept of hair care and skincare. Consistent with the findings in Study 3, makeup products, even in the case of concept testing, were more associated with social-oriented emotions such as ‘fun’ and ‘adult/professional’. They were also more related to emotions such as ‘glamorous’ and ‘pretty/beautiful/attractive’. This could be explained by the fact that people use makeup products for vanity purposes (Talavera & Sasse, 2019). Positive emotions associated with the skincare and hair care concepts were more about sensory pleasure and self-wellbeing. As they had different functions, dry shampoo was more likely to elicit positive emotions such as ‘neat’, ‘ready’, ‘approachable’, and ‘comfortable’, while the avocado mask made consumers feel more ‘pampered’, ‘nourished’, ‘calm/soothed’ and ‘relaxed’.

The PCA biplot also suggested how the six discriminating negative emotions were associated to these concepts. Feeling ‘ridiculous’ was more likely to be a negative emotional response elicited by makeup products than skincare and hair care. ‘Unnatural’ was more associated with the two makeup concepts and the dry shampoo concepts. There were several negative emotions that might be more specific to describe the negative experiences related to the dry shampoos. They were ‘sloppy/messy/unkempt’, ‘bland/dull’, ‘insecure/unconfident’. These emotions could be elicited by consumers’ uncertainty about the effect and the function of this product as people only use this when they don’t wash their hair. As the four product concepts were all well liked, these negative emotions were all used by less than 10% of the participants to describe the four concepts, except for one emotion-‘sloppy/messy/unkempt’ that was used by 14% of the participants, and the mean intensities of these emotions were all very low.

Table 2.8. Mean ratings of the significant emotions across the four product concepts

Emotion Terms	**Avocado Mask**	**Dry Shampoo**	**Lip Balm Oil**	**Mascara**
Adult/Professional (+)[1]	0.23	0.22	0.44	0.41

Approachable (+)	0.14 b	0.55 a	0.43 a	0.49 a
Awake (+)	0.51 a	0.25 b	0.51 a	0.17 b
Awesome/Wonderful (+)	0.24 b	0.16 b	0.61 a	0.49 a
Balanced (+)	0.63 a	0.38 b	0.29 b	0.17 b
Bland/Dull (-)	0.01 b	0.25 a	0.03 b	0.01 b
Calm/Soothed (+)	1.66 a	0.4 b	0.15 b	0.23 b
Clean (+)	1.45 a	1.46 a	0.66 b	0.47 b
Comfortable (+)	0.49 b	0.77 a	0.62 ab	0.41 b
Confident (+)	0.43 c	0.81 b	1.26 a	1.13 ab
Fresh (+)	1.97 a	1.63 a	1.01 b	0.87 b
Fun (+)	0.15 c	0.17 c	0.72 b	1.2 a
Glamorous (+)	0.28 b	0.27 b	1.06 a	1.25 a
Glowing (+)	1.21 b	0.23 c	0.47 c	1.56 a
Happy (+)	0.29 b	0.55 ab	0.58 a	0.79 a
Healthy (+)	1.79 a	0.71 b	0.85 b	0.59 b
Incomplete (-)	0.05 ab	0.15 a	0.03 b	0.01 b
Insecure/Unconfident (-)	0.01 b	0.17 a	0.03 b	0.06 ab
Luxurious (+)	0.7 a	0.16 b	0.59 a	0.73 a
Natural (+)	1.27 a	1.17 ab	1.19 a	0.83 b
Neat (+)	0.17 c	0.58 a	0.4 ab	0.22 bc
Nourished (+)	2.04 a	0.65 c	0.64 c	1.15 b
Pampered (+)	1.61 a	0.33 c	0.5 bc	0.79 b
Perfect/Flawless (+)	0.39 a	0.18 b	0.37 ab	0.35 ab
Pretty/Beautiful/Attractive (+)	0.8 b	0.49 b	1.37 a	1.58 a
Put-together/Polished (+)	0.66 b	0.98 b	1.35 a	1.59 a

Ready (+)	0.34 c	1.51 a	0.89 b	0.89 b
Refreshed/Rejuvenated/Energized (+)	1.55 a	1.1 b	0.38 c	0.52 c
Relaxed (+)	1.11 a	0.3 b	0.14 b	0.11 b
Relieved (+)	0.29 a	0.29 a	0.05 b	0.01 b
Ridiculous (-)	0.02 b	0.02 b	0.09 ab	0.15 a
Sexy (+)	0.04 c	0.11 c	0.49 b	0.92 a
Sloppy/Messy/Unkempt (-)	0.07 b	0.39 a	0.09 b	0.08 b
Sophisticated (+)	0.18 b	0.13 b	0.62 a	0.46 a
Special (+)	0.33 a	0.07 b	0.43 a	0.44 a
Unnatural (-)	0.04 b	0.15 ab	0.05 ab	0.21 a
Vibrant (+)	0.59 b	0.17 c	0.66 ab	0.9 a
Youthful (+)	1.23 a	0.21 d	0.54 c	0.87 b

[1]. This term was found to be significant by ANOVA but non-significant in post-hoc tests. Means with the same letter designation within each row were not statistically different (alpha=0.05).

Table 2.9. Percentages of selection for significant emotions across the four product concepts

Emotion Terms	Avocado Mask	Dry Shampoo	Lip Balm Oil	Mascara
Adult/Professional[1]	9%	9%	2%	2%
Approachable (+)	6% b	23% a	21% a	17% a
Awake (+)	20% a	11% ab	8% b	19% a
Awesome/Wonderful (+)	9% bc	7% c	18% ab	23% a
Balanced (+)	25% a	17% ab	7% b	12% b
Bland/Dull (-)	1% b	9% a	1% b	1% b
Calm/Soothed (+)	61% a	17% b	11% b	7% b
Clean (+)	53% a	56% a	19% b	25% b
Comfortable (+)	21% b	35% a	17% b	25% ab

Confident (+)	17% b	32% a	45% ab	46% a
Fresh (+)	71% a	63% a	32% b	39% b
Fun (+)	6% c	7% c	46% a	29% b
Glamorous (+)	11% b	11% b	47% a	39% a
Glowing (+)	44% a	9% b	57% a	19% b
Happy (+)	11% b	23% a	31% a	22% ab
Healthy (+)	67% a	27% b	23% b	31% b
Incomplete (-)	2% ab	7% a	1% b	1% b
Insecure/Unconfident (-)	1% b	7% a	3% ab	1% b
Luxurious (+)	28% a	7% b	29% a	24% a
Natural (+)	47% a	43% ab	31% b	44% ab
Neat (+)	7% c	23% a	9% bc	17% ab
Nourished (+)	73% a	27% c	43% b	25% c
Pampered (+)	59% a	14% b	33% a	20% ab
Pretty/Beautiful/Attractive (+)	33% b	19% c	59% a	49% a
Put-together/Polished (+)	25% c	40% b	58% a	49% ab
Ready (+)	15% c	57% a	35% b	33% b
Refreshed/Rejuvenated/Energized (+)	55% a	43% a	21% b	15% b
Relaxed (+)	41% a	13% b	5% b	7% b
Relieved (+)	11% a	13% a	1% b	2% b
Ridiculous (-)	1% ab	1% b	6% a	4% ab
Sexy (+)	1% c	5% c	33% a	20% b
Sloppy/Messy/Unkempt (-)	3% b	17% a	3% b	4% b
Sophisticated (+)	8% bc	5% c	17% ab	24% a
Special (+)	13% a	3% b	18% a	17% a

Unnatural (-)[1]	2% a	8% a	8% a	2% a
Vibrant (+)	23% b	7% c	37% a	27% ab
Youthful (+)	49% a	9% c	34% b	22% b

[1.] This term was found to be significant by Cochran's Q tests but non-significant in post-hoc tests.
Percentages with the same letter designation within each row were not statistically different (alpha=0.05).

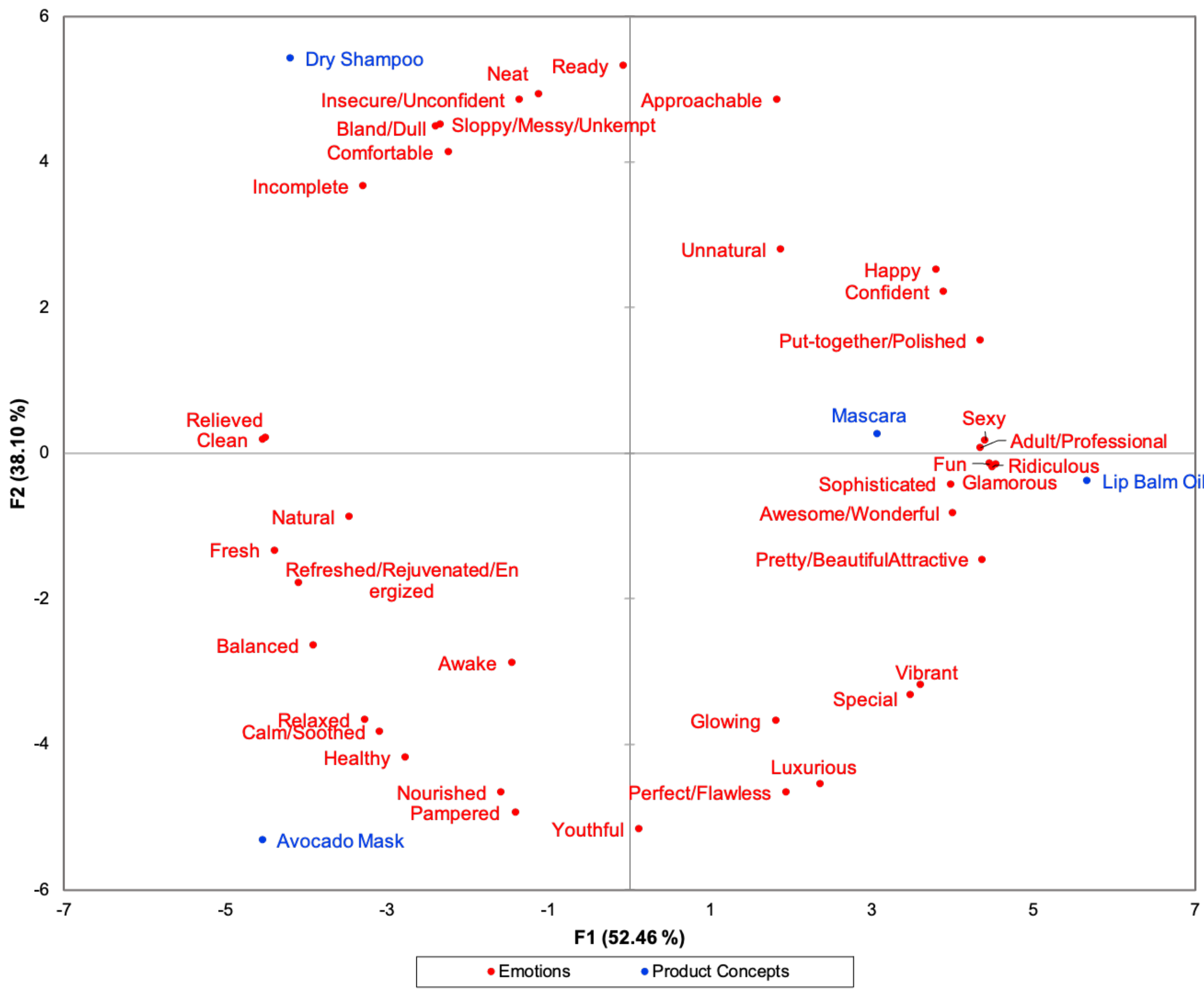
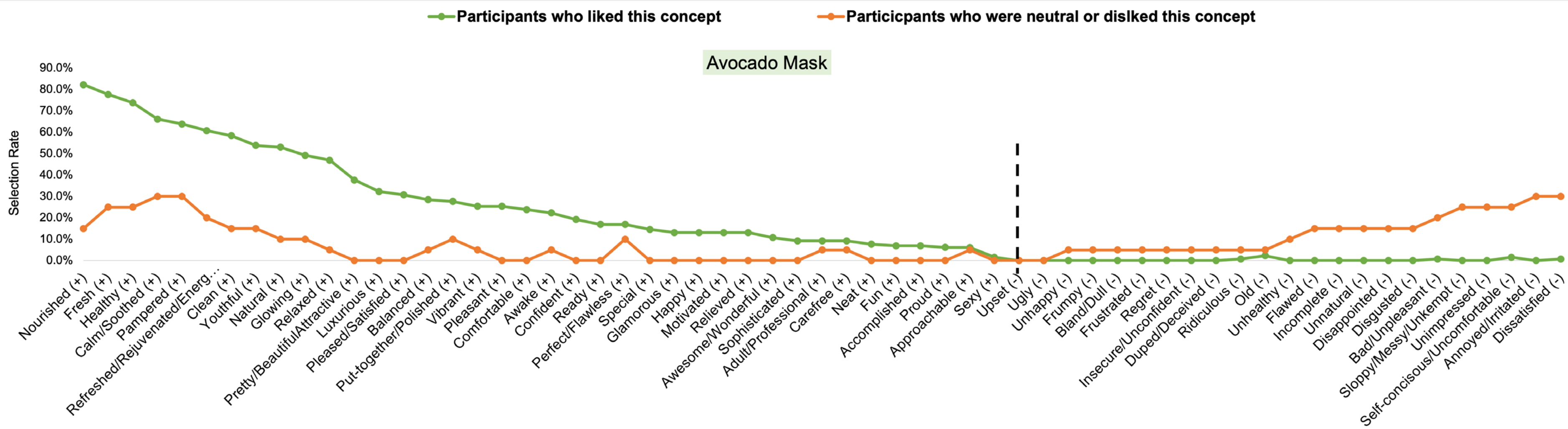


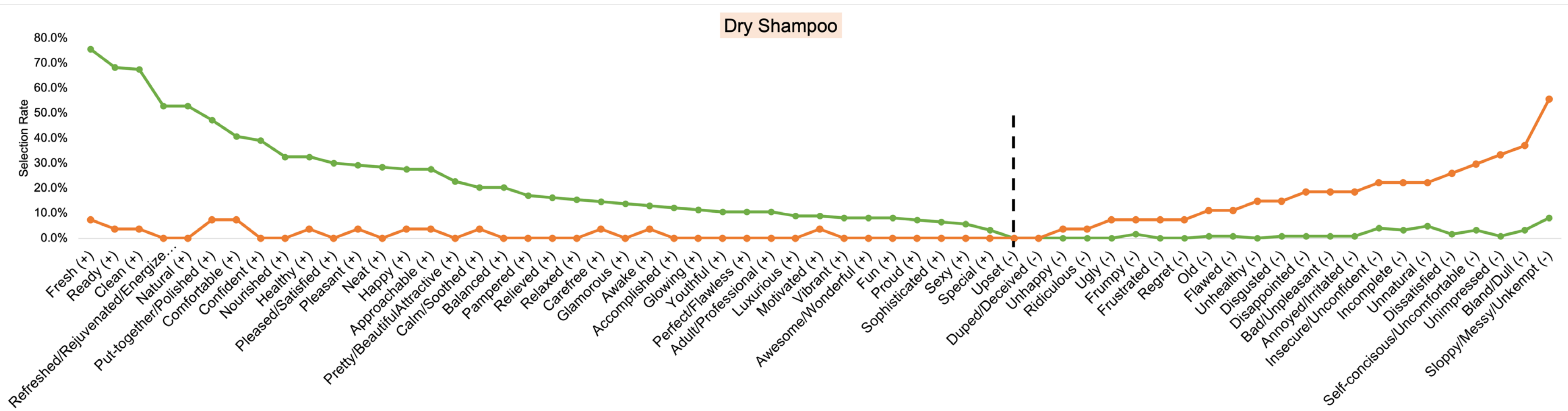
Figure 2.5. First two dimensions of PCA on the mean RATA intensities of significant emotions for the four product concepts

The Final Emotion Lexicon

As all product concepts were well-liked in the test, the usage rate of most negative emotion was low when it was calculated based on the results of all participants. This made it hard to examine which negative emotions were more relevant with the beauty care category. For validation purposes, within each product, the usage/selection percentage of each emotion was adjusted by considering only a subgroup of participants. We first divided participants into two groups within each product concept: those who liked the product concept (overall liking >5) and those who had neutral opinions and disliked the concept (overall liking <= 5). Then the percentage of selection for each positive emotion term was calculated for each product only considering participants who liked the concept. The same calculation was done for negative emotion terms for each product only considering participants who were neutral or disliked the product concept. The adjusted percentage of selection by each consumer group within each product can be seen in Figure 2.6. If any positive terms that were used by less than 10% of people who liked a concept, or negative terms were used by less than 10% of people who were neutral or dislike a concept for all the four product concepts, then these terms were removed from the initial lexicon. This step led to the removal of four terms: proud, upset, ugly and unhappy. The final emotion lexicon contains a total of 57 emotion terms with 37 positive emotions and 20 negative emotions (Table 2.10).

Based on the adjusted percentage of selection for each emotion in this study and the usage frequency data in Study 3, recommendations on how to use this lexicon for different categories of beauty care were provided: if any positive emotion terms were used by more than 20% of the participants in survey 3 to describe their favorite product of a specific category (skincare, hair care and makeup), or used by more than 20% participants who liked a product concept of that category in study 4, then those terms were considered applicable for that beauty category; if any negative emotion terms were used by more than 15% of the participants in survey 3 to describe their least favorite product of a specific category (skincare, hair care and makeup), or used by more than 15% participants who disliked a product concept of that category in study 4, then those terms were considered applicable for that beauty category. These results were shown in Table 2.10. The positive and negative emotions recommended to the three categories of beauty care were not mutually exclusive. Of the 37 positive emotion terms in the lexicon, 26, 27 and 29 terms were identified to be applicable to skincare, hair care and makeup respectively. 17, 16, 16 out of the 20 negative terms were found to be applicable to skincare, hair care and makeup respectively.


Participants who liked this concept
Particicpants who were neutral or dislked this concept
Avocado Mask
Selection Rate


Dry Shampoo
Selection Rate

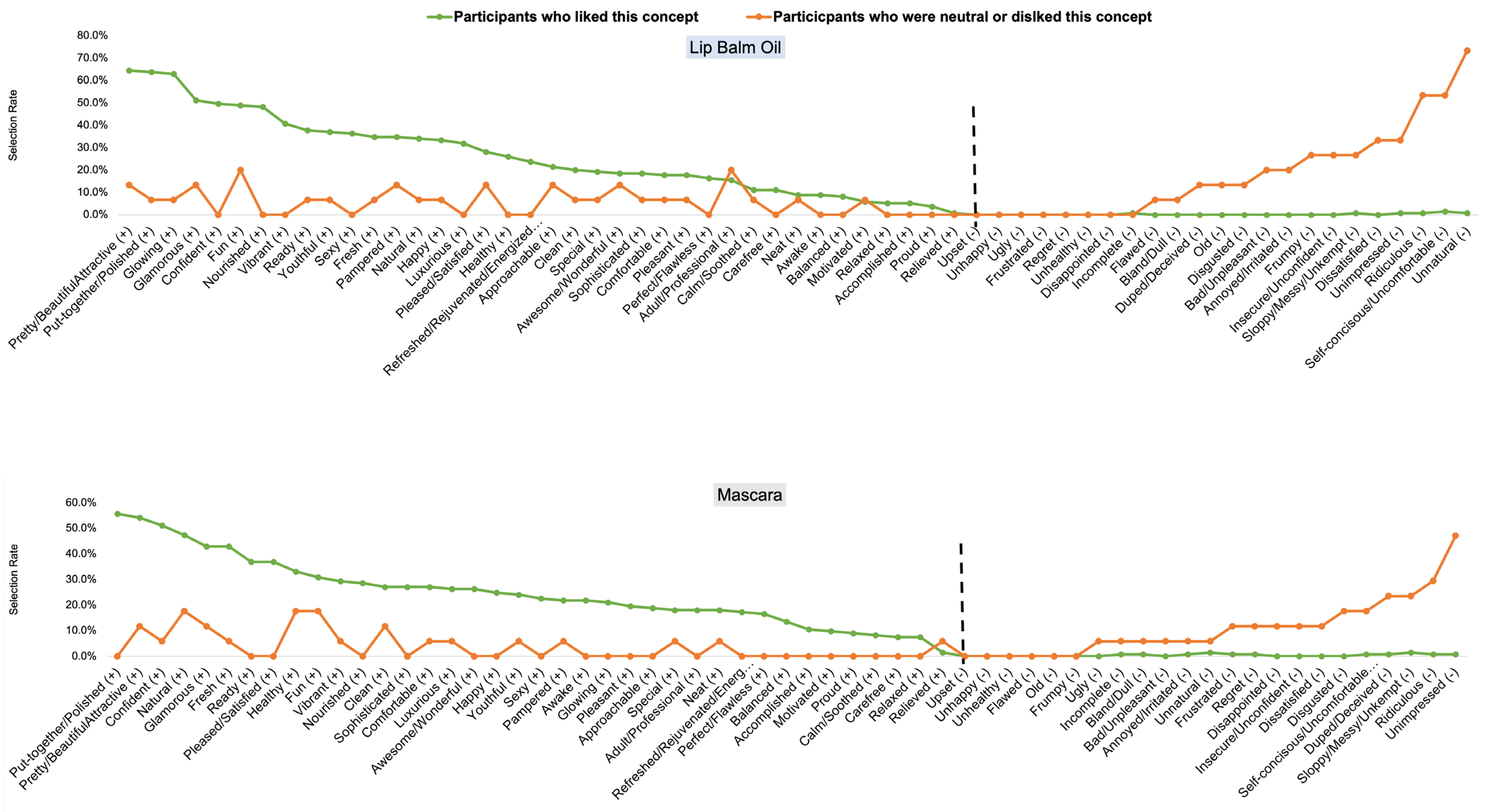


Figure 2.6. Adjusted selection rate for each emotion in each product concept

Table 2.10. Final emotion lexicon for beauty care products

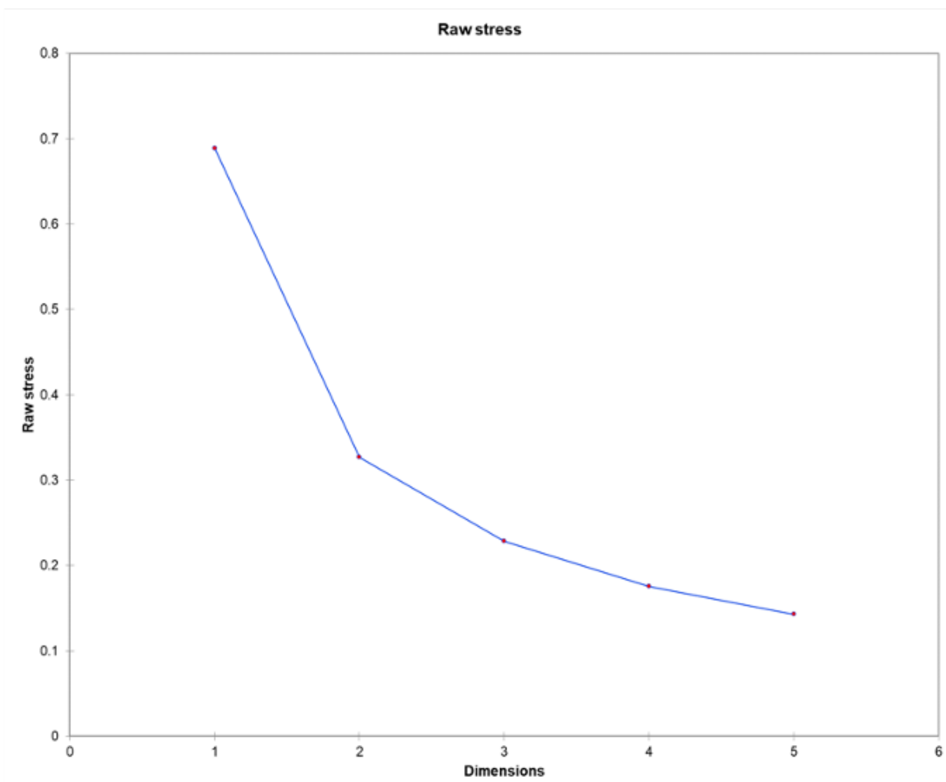
Positive Emotions (n=37)	Skincare (n=26)	Hair Care (n=27)	Makeup (n=29)
Clean	✓	✓	✓
Comfortable	✓	✓	✓
Confident	✓	✓	✓
Fresh	✓	✓	✓
Happy	✓	✓	✓
Healthy	✓	✓	✓
Luxurious	✓	✓	✓
Natural	✓	✓	✓
Nourished	✓	✓	✓
Pampered	✓	✓	✓
Pleased/Satisfied	✓	✓	✓
Pretty/Beautiful/Attractive	✓	✓	✓
Put-together/Polished	✓	✓	✓
Ready	✓	✓	✓
Refreshed/Rejuvenated/Energized	✓	✓	✓
Vibrant	✓	✓	✓
Balanced	✓	✓	
Calm/Soothed	✓	✓	
Neat	✓	✓	
Pleasant	✓	✓	
Relaxed	✓	✓	
Awake	✓		✓
Glowing	✓		✓
Perfect/Flawless	✓		✓
Youthful	✓		✓
Approachable		✓	✓
Awesome/Wonderful		✓	✓
Glamorous		✓	✓
Sexy		✓	✓
Relieved	✓		
Accomplished		✓	
Carefree		✓	
Adult/Professional			✓
Fun			✓
Motivated			✓
Sophisticated			✓
Special			✓

Negative Emotions (n=20)	Skincare (n=17)	Hair Care (n=16)	Makeup (n=16)
Annoyed/Irritated	✓	✓	✓
Bad/Unpleasant	✓	✓	✓
Disappointed	✓	✓	✓
Disgusted	✓	✓	✓
Dissatisfied	✓	✓	✓
Flawed	✓	✓	✓
Frustrated	✓	✓	✓
Insecure/Unconfident	✓	✓	✓
Self-conscious/Uncomfortable	✓	✓	✓
Sloppy/Messy/Unkempt	✓	✓	✓
Unimpressed	✓	✓	✓
Unnatural	✓	✓	✓
Bland/Dull	✓	✓	
Incomplete	✓	✓	
Regret	✓	✓	
Unhealthy	✓	✓	
Duped/Deceived	✓		✓
Frumpy			✓
Old			✓
Ridiculous			✓

Structure of The Final Emotion Lexicon

The structure of the emotion lexicon was explored by MDS on correlation matrix (Figure 2.7). The scree plot of Kruskal's raw stress indicated a 2-dimensional solution was the most appropriate. A stress value of 0.3 is not an excellent representation of the proximities of the emotions in the lexicon, however, the larger patterns between the terms should be visible (Borg et al., 2012). The distribution of emotions on the MDS map seemed to partially follow the bi-dimensional (valence-arousal) models of emotions (Russell et al., 1989). The first dimension clearly differentiated positive and negative emotions. However, the arousal/activation pattern was only found in the second dimension for positive emotions. There might be more dimensions existing in emotions related to beauty care.

As the raw stress of the two-dimensional MDS was relatively large, the relationships between emotions, especially between those with short distances on the MDS map, might not be represented accurately because of the large distortions. Consequently, we performed additional cluster analysis on the same correlation matrix to further examine the relationships between emotions in the lexicon. The AHC dendrogram (Figure 2.8) was truncated at 14 main clusters by minimizing variation and maximizing semantic similarity within each cluster. The 14 clusters were summarized in Table 2.11 with 9 clusters for positive emotions and 5 clusters for negative emotions. It was noteworthy that emotions of the same cluster might not necessarily always be semantically similar, instead, they were emotions that always happened together in the context of beauty care products.



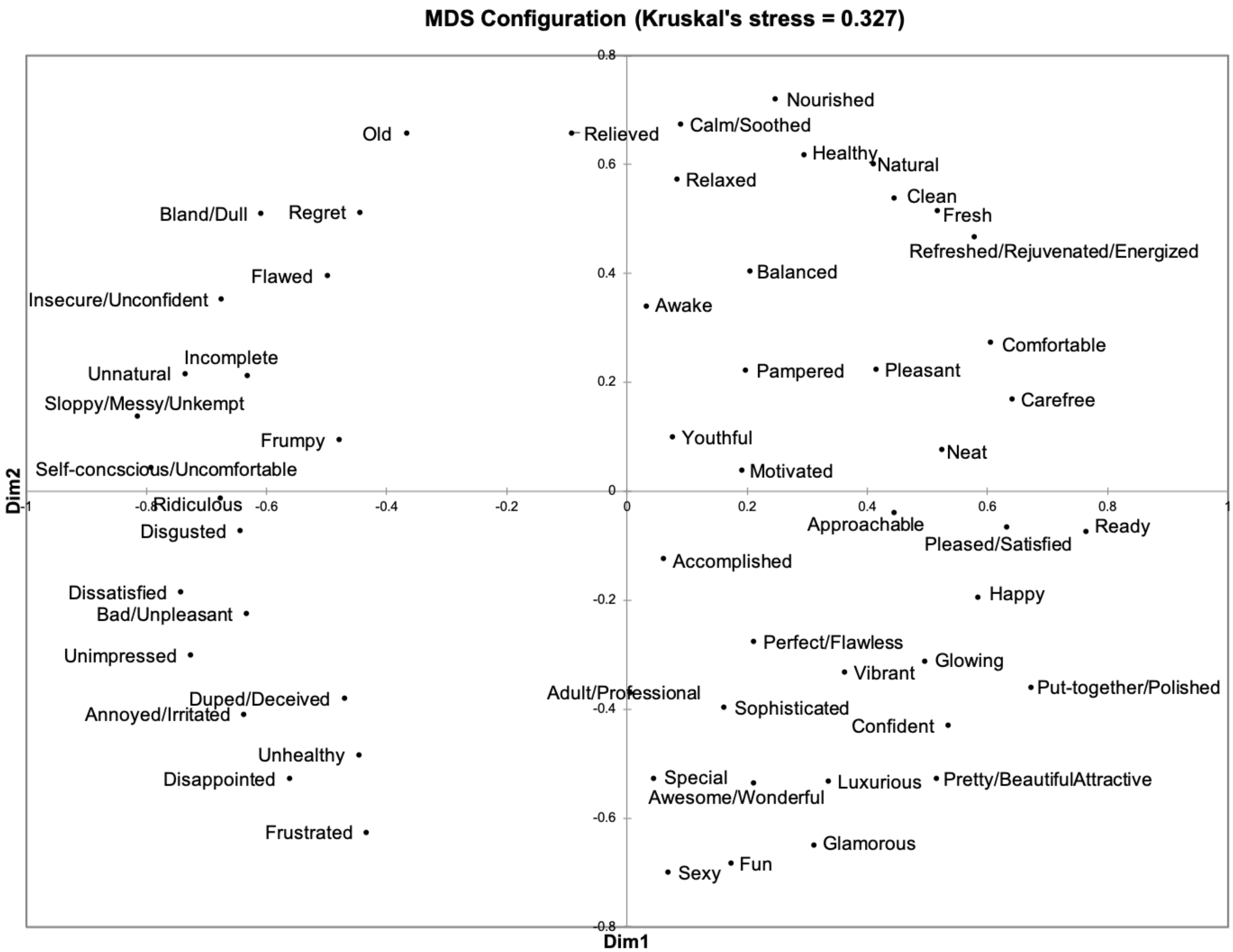


Figure 2.7. Two-dimensional representation of the emotion terms in the lexicon constructed by MDS on correlation matrix

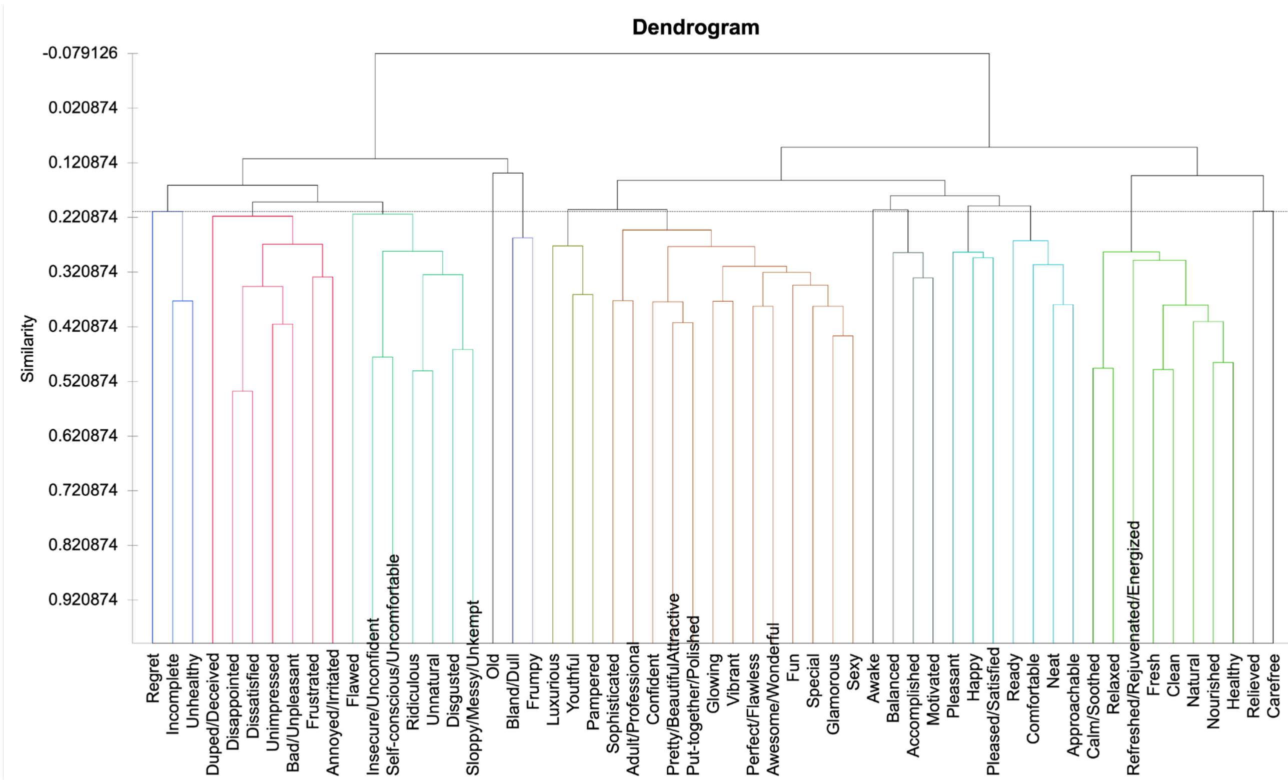


Figure 2.8. Dendrogram of AHC using correlation dissimilarity and unweighted pair-group average agglomeration method

Table 2.11. Clusters of emotions for beauty care products based on AHC

Cluster 1	**Cluster 2**	**Cluster 3**	**Cluster 4**
Confident	Calm/Soothed	Comfortable	Accomplished
Fun	Fresh	Neat	Motivated
Glamorous	Nourished	Ready	Balanced
Sophisticated	Refreshed/Rejuvenated/Energized	Approachable	
Adult/Professional	Healthy		
Perfect/Flawless	Clean		
Pretty/Beautiful/Attractive	Natural		
Put-together/Polished	Relaxed		
Sexy			
Awesome/Wonderful			
Glowing			
Special			
Vibrant			

Cluster 5	**Cluster 6**	**Cluster 7**	**Cluster 8**	**Cluster 9**

Happy	Luxurious	Relieved	Carefree	Awake
Pleasant	Youthful			
Pleased/Satisfied	Pampered			
Cluster 10	**Cluster 11**	**Cluster 12**	**Cluster 13**	**Cluster 14**
Disappointed	Ridiculous	Incomplete	Bland/Dull	Old
Dissatisfied	Flawed	Regret	Frumpy	
Frustrated	Disgusted	Unhealthy		
Unimpressed	Insecure/Unconfident			
Annoyed/Irritated	Unnatural			
Bad/Unpleasant	Self-conscious/Uncomfortable			
Duped/Deceived	Sloppy/Messy/Unkempt			

Overall Discussion

This study described a detailed procedure in four studies for the development of an emotion lexicon that can be used to profile consumers' emotional responses to different categories of beauty care products. The development process started with extracting relevant terms from consumers' active languages-online product reviews in Study 1, which resulted in a primary collection of 330 terms. The generation of this list of terms was achieved by filtering all words from online reviews through sentimental lexicons (Hu & Liu, 2004; Mohammad & Turney, 2010, 2013; Nielsen, 2011) and the manual selection by the researchers. It was important to note that not all terms were 'true' emotion terms, as the sentimental lexicons contain words that are not only emotions, but also any words associated with emotions/sentiments; for example, nouns such as beer, cash, music were included in the lexicon as related to positive valence in sentiment (Mohammad & Turney, 2010, 2013). This makes further manual examination necessary on the list of terms filtered through the sentimental lexicon. In manual selection of terms for the initial list, the researchers adopted a broad sense for the definition of emotion, which was any terms that can be used to describe subjective affective experience including emotions, moods, attitudes, or evaluative words. The adoption of a broad scope of affective feelings has been found common in developing emotion lexicons associated with consumer goods (Chrea et al., 2008; Gmuer et al., 2015; King & Meiselman, 2010; Spinelli et al., 2014; Thomson & Crocker, 2013). Both emotions and other affective feelings were of interest when understanding consumer experiences as product perception was mediated by not only emotions elicited in the moment of consumption, but also by feelings and abstract conceptualizations associated with products in the mind of consumers (Spinelli et al., 2014; Thomson et al., 2010). 233 terms from the primary collection of 330 terms were found to be reported in previous emotion lexicons or associated research (Chrea et al., 2008; King & Meiselman, 2010; Laros & Steenkamp, 2005; Talavera & Sasse, 2019). This suggests text analysis of online reviews are effective in identifying consumer terminology not only related to the sensory features of the products but also emotional profiles of the products (Hamilton & Lahne, 2020; Kim & Kang, 2018). More importantly, sourcing terms from the voice of consumers on a diversity of beauty care products has laid the foundation for developing a consumer-friendly emotion lexicon as terms would be more familiar, and less confusing in the context of beauty care to consumers, compared to a lexicon sourced from psychological literature (King & Meiselman, 2010b). Furthermore, we conducted fine-grained consumer surveys with follow-up interviews to ensure all terms included in the lexicon were the most relevant terms to the positive and negative experience of three main categories of beauty care products.

Product concepts were used as stimuli to validate the initial emotion lexicon developed from Study 1 to Study 3. The results showed that more than 80% of the positive emotions in the lexicon were able to differentiate the tested concepts from different product categories, while only 26% of negative terms were significant between the four concepts. This might be explained by two reasons. First, as study 3 suggested, positive experiences of different categories of beauty care tended to elicit different types of emotions, while negative experiences of skincare, hair care and makeup elicited similar negative emotions. Secondly, all product concepts were well-liked, resulting in low usage rates for most of the negative emotions. The high hedonic ratings for all four concepts were not surprising. Consumers have been found to rate product concepts based on their expected perception, as a result, they tend to be more positive when evaluating a concept compared to the real product in a blind test (Ares & Varela, 2018). Moreover, it should be noted that the emotion terms generated and validated in this lexicon may not only be the results of the sensory characteristics of beauty care products. This was because all the emotion terms were sourced from online reviews where consumers expressed their comments based on their whole product experiences. In addition, the list of emotions was reduced and validated in online surveys using either context (favorite or least favorite product) or product concepts as main stimuli, but no blind test was incorporated. As a result, our emotion lexicon contained emotions elicited by both intrinsic and extrinsic characteristics of product experiences, which enable its application in both blind sensory testing and concept testing. As for sources of emotions in the experiences of consumer goods, Desmet & Schifferstein. (2008) summarized positive and negative emotions in food experience came from direct conditions (sensory properties and experienced consequences) and indirect conditions (expectations and associations (anticipated consequences), personal or cultural meanings, and actions of associated agents).

A total of 37 positive emotions and 20 negative emotions have been selected to form the final lexicon for beauty care. Considering different types of beauty care might elicit different emotions, we provided further recommendations on how to use this lexicon to each category of skincare, hair care and makeup products. One may have concerns about the length of the lexicon when being applied to a commercial testing. In previous emotion lexicons for food, for example, researchers suggested testing fewer than three food samples when using the EsSense profile which contains a total of 39 terms (King et al., 2013). In our case, the primary goal was to include as many terms relevant to different categories of beauty care products as possible, even for negative emotions, to ensure no important dimensions are missing. This lexicon has been developed as a starting point/ a base lexicon that future researchers can use to generate their product-specific lexicons with fewer terms. As a result, the length was less an issue. Together with the lexicon was the structure of these emotion terms in MDS and AHC configurations. MDS suggested that the emotional space of beauty care might not be fully represented by a bidimensional model. This finding was in line with the those by Chrea et al. (2008) when researching odor-associated emotions. Cluster analysis (AHC) uncovered a detailed relationships between emotion terms based on correlation. Future users of this lexicon may not necessarily use the full lexicon but select relevant emotions from each cluster we provided to form a reduced list that can be applied to testing multiple products.

Conclusion

An emotion lexicon containing 37 positive emotions and 20 negative emotions has been developed for beauty care products and validated between product categories. Skincare and hair care were found to elicit emotions related to self-well-being and sensory pleasure, while makeup evoked more social-oriented emotions and emotions related to vanity. The emotion lexicon developed in this study provides a good starting point for the further development of product-specific emotions in the beauty care category, as well as other emotion research such as exploring sensory and emotion associations. Text analysis of online reviews representing the voice of consumers could facilitate sensory and consumer research, especially for terminology development.

There were limitations for this research. First, the emotion lexicon was developed considering only three major categories of beauty care products (skincare, hair care and makeup). Therefore, the terminology in the developed lexicon might not be able to cover all categories such as beauty tools or perfumes. Secondly, identification and refinement of the terms for the lexicon were based on consumer responses from online surveys using either verbal contexts or concepts as stimuli. These emotional responses to verbal contexts or concepts may depend more on memory or expectations rather than immediate product experience. As a result, this lexicon should be further tested for its ability to discriminate different sensory profiles in central location tests and home use tests. Finally, the emotion lexicon was only validated for its discriminating ability within product categories. Future research should examine its performance in a within product-category design.

References

Apaolaza-Ibáñez, V., Hartmann, P., Diehl, S., & Terlutter, R. (2011). Women satisfaction with cosmetic brands: The role of dissatisfaction and hedonic brand benefits. 11.

Ares, G., Bruzzone, F., Vidal, L., Cadena, R. S., Giménez, A., Pineau, B., Hunter, D. C., Paisley, A. G., & Jaeger, S. R. (2014). Evaluation of a rating-based variant of check-all-that-apply questions: Rate-all-that-apply (RATA). Food Quality and Preference, 36, 87–95.

Ares, G., & Varela, P. (Eds.). (2018). Methods in consumer research. WP, Woodhead Publishing, an imprint of Elsevier.

Bhumiratana, N., Adhikari, K., & Chambers, E. (2014). The development of an emotion lexicon for the coffee drinking experience. Food Research International, 61, 83–92.

Borg, I., & Lingoes, J. (2012). Multidimensional similarity structure analysis. Springer Science & Business Media.

Chaya, C., Eaton, C., Hewson, L., Vázquez, R. F., Fernández-Ruiz, V., Smart, K. A., & Hort, J. (2015). Developing a reduced consumer-led lexicon to measure emotional response to beer. Food Quality and Preference, 45, 100–112.

Chrea, C., Grandjean, D., Delplanque, S., Cayeux, I., Le Calve, B., Aymard, L., Velazco, M. I., Sander, D., & Scherer, K. R. (2008). Mapping the Semantic Space for the Subjective Experience of Emotional Responses to Odors. Chemical Senses, 34(1), 49–62.

David, O. A., David, D. O., Mogoase, C., Popescu, L. C., Giosan, C., & Pellegrino, A. (2019). Psychological effects and brain correlates of a rose-based scented cosmetic cream. Journal of Sensory Studies, 34(6), e12536.

Desmet, P. M. A., & Schifferstein, H. N. J. (2008). Sources of positive and negative emotions in food experience. Appetite, 50(2–3), 290–301.

Ferdenzi, C., Delplanque, S., Barbosa, P., Court, K., Guinard, J.-X., Guo, T., Craig Roberts, S., Schirmer, A., Porcherot, C., Cayeux, I., Sander, D., & Grandjean, D. (2013). Affective semantic space of scents. Towards a universal scale to measure self-reported odor-related feelings. Food Quality and Preference, 30(2), 128–138.

Gmuer, A., Nuessli Guth, J., Runte, M., & Siegrist, M. (2015). From emotion to language: Application of a systematic, linguistic-based approach to design a food-associated emotion lexicon. Food Quality and Preference, 40, 77–86.

Guest, S., Dessirier, J. M., Mehrabyan, A., McGlone, F., Essick, G., Gescheider, G., Fontana, A., Xiong, R., Ackerley, R., & Blot, K. (2011). The development and validation of sensory and emotional scales of touch perception. Attention, Perception, & Psychophysics, 73(2), 531–550.

Hamilton, L. M., & Lahne, J. (2020). Fast and automated sensory analysis: Using natural language processing for descriptive lexicon development. Food Quality and Preference, 83, 103926.

Hu, M., & Liu, B. (2004). Mining and Summarizing Customer Reviews. 10.

Kim, S. G., & Kang, J. (2018). Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews. Information Processing & Management, 54(6), 938–957.

Kao Corporation | Research on Skin Care with a Focus on Stimulation that Creates Pleasant Feelings. (2018). Retrieved January 5, 2021, from https://www.kao.com/global/en/news/rd/2018/20181115-001/

King, S. C., & Meiselman, H. L. (2010). Development of a method to measure consumer emotions associated with foods. Food Quality and Preference, 21(2), 168–177.

King, S. C., Meiselman, H. L., & Thomas Carr, B. (2013). Measuring emotions associated with foods: Important elements of questionnaire and test design. Food Quality and Preference, 28(1), 8–16.

Lagast, S., Gellynck, X., Schouteten, J. J., De Herdt, V., & De Steur, H. (2017). Consumers' emotions elicited by food: A systematic review of explicit and implicit methods. Trends in Food Science & Technology, 69, 172–189.

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.

Meiselman, H. L. (2016). Emotion Measurement. In Emotion Measurement (pp. 645–697). Elsevier.

Meyners, M., Jaeger, S. R., & Ares, G. (2016). On the analysis of Rate-All-That-Apply (RATA) data. Food Quality and Preference, 49, 1–10.

Mohammad, S. M., & Turney, P. D. (2010). Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon.

Mohammad, S. M., & Turney, P. D. (2013). CROWDSOURCING A WORD-EMOTION ASSOCIATION LEXICON. Computational Intelligence, 29(3), 436–465.

Ng, M., Chaya, C., & Hort, J. (2013). The influence of sensory and packaging cues on both liking and emotional, abstract and functional conceptualisations. Food Quality and Preference, 29(2), 146–156.

Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. ArXiv:1103.2903 [Cs].

Painchault, T., Perrin, L., Loijens, L. W. S., Linden, X. van der, Theuws, H. M. M., & Zimmerman, P. H. (2020). Application of scientific measures to demonstrate the relaxing properties of Peony fragrance in hair care products. Journal of Sensory Studies, 35(2), e12560.

Parente, M. E., Manzoni, A. V., & Ares, G. (2011). EXTERNAL PREFERENCE MAPPING OF COMMERCIAL ANTIAGING CREAMS BASED ON CONSUMERS' RESPONSES TO A CHECK-ALL-THAT-APPLY QUESTION: PREFERENCE MAPPING OF ANTIAGING CREAMS BASED ON A CATA QUESTION. Journal of Sensory Studies, 26(2), 158–166.

Porcherot, C., Delplanque, S., Raviot-Derrien, S., Calvé, B. L., Chrea, C., Gaudreau, N., & Cayeux, I. (2010). How do you feel when you smell this? Optimization of a verbal measurement of odor-elicited emotions. Food Quality and Preference, 21(8), 938–947.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A Single-Item Scale of Pleasure and Arousal. 10.

Silge, J., & Robinson, D. (2016). tidytext: Text Mining and Analysis Using Tidy Data Principles in R. The Journal of Open Source Software, 1(3), 37.

Spinelli, S., Masi, C., Dinnella, C., Zoboli, G. P., & Monteleone, E. (2014). How does it make you feel? A new approach to measuring emotions in food product experience. Food Quality and Preference, 37, 109–122.

Talavera, M., & Sasse, A. M. (2019). Gathering consumer terminology using focus groups—An example with beauty care. Journal of Sensory Studies, 34(6).

Thomson, D. M. H., & Crocker, C. (2013). A data-driven classification of feelings. Food Quality and Preference, 27(2), 137–152.

Thomson, D. M. H., Crocker, C., & Marketo, C. G. (2010). Linking sensory characteristics to emotions: An example using dark chocolate. Food Quality and Preference, 21(8), 1117–1125.

Chapter 3 - Exploring Sensory Drivers of Liking and Emotional Associations for Beauty Care: A Case Study with Hand Creams

Abstract

In beauty product development, it is key to understand which sensory characteristics of the product drive consumer experience. The objective of this study was to identify the main sensory drivers of liking and emotions for hand creams in different consumer segments. Twelve hand cream samples were evaluated by a trained descriptive panel for aroma, appearance, and texture & skinfeel (pick-up, rub-out and afterfeel). Seven hand creams selected from the descriptive sensory space were rated for overall liking, emotions (using the lexicon developed from chapter 2), and consumer perception (attribute intensity and CATA) in a home use test (HUT) with one hundred female consumers. Cluster analysis and external preference mapping identified three consumer clusters with different liking patterns: the thick & waxy texture likers, mild scent & low-medium thickness likers, and strong scent likers. Consumers with different liking patterns differed in their emotional associations with sensory characteristics of hand creams. However, high intensities of certain aroma attributes seemed to elicit high-arousal emotions for all groups. Comparing consumer characterization using CATA with descriptive analysis configuration helped interpret these differences between consumer segments.

Overall, waxy texture elicited emotions including fresh, vibrant, and refreshed/rejuvenated/energized within the thick & waxy liker group. This group also characterized the waxy & thick as feeling rich. Wet & spreadable texture and oily afterfeel were described as lacking efficacy for this group, and they elicited negative emotions such as duped/deceived and flawed. The mild scent & low-medium thickness likers related low overall strength of aroma and low sweet aromatic with feeling natural. Thick texture was negatively

associated with feeling clean, fresh, refresh /rejuvenated/energized. They described the thin and less thick texture as more hydrating to their hands. The strong scent likers associated high overall aroma strength positively with awake, nourished, pretty/beautiful/attractive and vibrant; and negatively with bland. Different aroma attributes elicited diverse emotions for this group. In general, the differences in emotion associations across consumer clusters seems to be due to the values and efficacy each cluster attached to the sensory characteristics of the hand creams. The findings of this study could guide the development of new hand creams products targeting at different consumer segments.

Introduction

In a fast-moving industry flooded with new products every year, cosmetic companies need to constantly offer discriminating products that are adapted to the needs and preferences of consumers to stay competitive in the marketplace. In addition to traditional consumer hedonic research, consumer responses beyond liking, such product-elicited emotions, are essential in understanding consumer experience as beauty care products are developed to make consumers feel better about themselves. Tactile and olfactory experience delivered by personal care products have been reported to impact consumers' emotions and moods (Chrea et al., 2008; David et al., 2019; Guest et al., 2011; Porcherot et al., 2010). As a result, emotional benefits of beauty care products are always conveyed to consumers via their sensory characteristics. For example, Painchault et al. (2020) claimed the peony fragrance added in shampoo could provide relaxing effect. Rose-scented cream was found to lead to more relaxation, satisfaction, happiness, and less stress compared to non-scented cream (David et al., 2019). Efficacy attributes of skin care such as absorbency may evoke emotions related to consumer trust and belief such as confident, secure, and happy (Sanderson & Hollowood, 2017).

Sensory drivers of liking and emotions for products in the beauty care category have been investigated by several researchers. For example, Xing et al. (2020) used design of experiment (DOE) to identify which combination of product characteristics would maximize consumer acceptance for makeup remover wipes. Parente et al. (2011) conducted external preference mapping for commercial antiaging creams based on consumer responses to CATA questions. The study suggested hydration, texture and perfume were the key drivers of liking, and more liked products evoked more positive emotional responses in consumers (Parente et al., 2011). However, these studies all used consumer perception and liking data as the only source to identify the key sensory drivers of liking and to study product-elicited emotions. Very limited research has used sensory data collected from trained panel to identify the sensory attributes that drive consumer liking and emotions for beauty care products, which is crucial for product development purpose. Furthermore, none of these studies on beauty care have considered individual differences (e.g., in liking, personal traits, etc.) when examining product-elicited emotions, which has been found important in food research (Bhumiratana et al., 2014; Mora et al., 2020; Schouteten et al., 2018). Sensory drivers of emotions were found to be moderated by consumer liking patterns in various categories of food products (Pierguidi et al., 2020; Spinelli et al., 2019).

This study aimed to fill this research gap in understanding the sensory drivers of liking and emotions for beauty care considering individual differences in liking patterns. In Chapter 2, a lexicon was developed to measure the product-elicited emotions for beauty care products (skincare, haircare, and makeup). Twenty-seven positive terms and 15 negative terms from this lexicon were identified as applicable to the skincare category. We used the lexicon developed for skincare for emotion measurement and hand creams as the testing samples in this study. The specific objectives were to 1) identify consumer segmentation of different liking patterns and sensory drivers of liking for each segment 2) explore sensory and emotional associations for different consumer segments; 3) interpret the differences in emotional association between consumer segments. The results of this study may provide valuable information for cosmetic companies in the development and marketing of their new products. In addition, this study demonstrated a workflow for the research of sensory and consumer behavior of beauty care products in product development setting.

Materials and Methods

Sensory Descriptive Analysis

Samples

Twelve hand cream products representing a wide range of brands, prices, claims, and sensory profiles were selected for the descriptive analysis study. The selection of hand cream samples with different sensory profiles was based on a perceptual map constructed from a projective mapping exercise on a total of twenty hand cream products with a group of 5 semi-trained panelists. Table 3.1 shows the detailed information about the 12 hand cream products selected for the descriptive analysis study. All samples were purchased either from local stores or Amazon and were stored at room temperature (21 ± 1 °C).

Table 3.1. Hand cream samples used in descriptive analysis

Product	Description
Dionis Vanilla Bean Goat Milk Hand Cream (Dionis, Warrington, PA)	Soothes and moisturizes, for dry hands; paraben free, cruelty-free
O'Keeffe's Working Hands Hand Cream (O'Keeffe's Company, Cincinnati, OH)	For extremely dry, cracked hands
EOS Shea Better Coconut Hand Cream (EOS Products, LLC., New York, NY)	24-hour hydration, paraben & phthalate-free, not tested on animals
Mrs. Meyer's Lemon Verbena Hand Lotion (The Caldrea Company, Racine, WI)	Paraben, phthalate and artificial colors-free
Love Beauty and Planet Coconut Milk & White Jasmine Hand Cream (Unilever, Trumbull, CT)	Paraben, dye-free, vegan, clean
Neutrogena Hydro Boost Gel Hand Cream with Hyaluronic Acid (Johnson & Johnson Consumer INC., Skillman, NJ)	Instantly boosts hydration for soft and supple hands
Shiseido Hand Cream-No Scented (Shiseido, Tokyo, Japan)	Prevents hand roughness
Burt's Bees Almond & Milk Hand Cream (Burt Bee's, INC., Durham, NC)	Keeps hands soft and smooth with emollient almond cream; 99% natural; paraben & phthalate-free
Gold Bond Ultimate Healing Hand Cream (Chattem, INC., Chattanooga, TN)	Repairs dry, problem hands; last through hand washing
Aveeno Skin Relief Hand Cream- Fragrance Free (Johnson & Johnson Consumer INC., Skillman, NJ)	Soothes very dry skin; intense moisture

Herbacin Kamille Hand Cream (Herbacin Cosmetic GmbH, Wutha-Farnroda, Germany)	Leaves rough, dry, chapped hands smooth and hydrated for a long time; no parabens, mineral oils, animal ingredients and synthetic colors
Nivea Crème Body, Face & Hand Moisturizing Cream (Beiersdorf INC., Wilton, CT)	Ideal for daily use for all intensive moisturizing needs

Descriptive Panel

Six highly trained panelists from the Center for Sensory Analysis and Consumer Behavior at Kansas State University, Manhattan, KS participated in the descriptive analysis. All panelists have received more than 1200 hours of general training on descriptive analysis including acuity for odors and texture, and descriptive techniques such as identifying, describing, ranking, or scoring a product attribute. In addition, the panelists had over 1000 hours of descriptive sensory experience, including non-food products such as skin creams, lotions, toothpaste, and soaps. The panel was conducted in December (Winter) 2020. As the study was conducted during Covid-19 pandemic, panelists chose the ways they felt most comfortable performing the descriptive study. Of the six panelists, one panelist participated in all sessions remotely via Zoom; testing samples, reference samples and testing materials were delivered to the panelist 30 minutes before each session. The other five panelists attended all sessions in person following the recommended social distancing protocols in the sensory testing room at Kansas State University. Products were evaluated by consensus.

Sample Preparation

Samples were prepared one hour before each evaluation session. All samples were prepared in containers labeled with random 3-digit blinding codes. Aroma attributes were evaluated in separate sessions from the evaluation of appearance, texture and skinfeel (pick-up, rub-out, afterfeel). For aroma evaluation, 5 grams of each sample were weighed into a 12 oz Styrofoam drinking cup with a lid (Dart Container Corporation, Mason, MI). The evaluation of appearance and texture involved multiple stages, as a result, each sample was prepared in different containers for attribute evaluation at each stage. For appearance evaluation, 0.1 ml of each sample was filled in a 1ml syringe (Henke Sass Wolf, Tuttlingen, Germany). Delivering samples from syringes helps keep the shape and size of samples consistent with each other for appearance evaluation. As some samples such as the O'Keefle's and Love Beauty and Planet were thick, samples used for texture and afterfeel evaluation were weighed in a tightly controlled amount into cups rather than being delivered by syringes. Samples for pick-up evaluation were weighed with 0.1 g of each in a 3.25 oz Soufflés plastic portion cup (Dart Container Corporation, Mason, MI) covered with a lid. Two sets of 0.05 g of each sample were weighed in 3.25 oz Soufflés plastic portion cups (Dart Container Corporation, Mason, MI) covered with lids for rub-out and afterfeel evaluation.

Aroma Evaluation

Panelists evaluated the aroma of samples neat, which meant all samples were presented in the Styrofoam cups without further manipulation (ASTM Standard E1490-19, 2019). Consensus method was used for attribute development and evaluation (Chambers, 2018). One 90-min orientation session was completed to introduce and familiarize the panel with the samples. During the orientation session, individual panelists first identified aroma attributes in each sample. Then after seeing all twelve samples to be tested, a list of aroma attributes, definitions and references were developed for these samples through consensus among panelists. A scale ranging from 0-15 with 0.5 increments (0= not present, 15= extremely high intensity) was used to rate the intensity of each aroma attribute. The twelve samples were evaluated in two days with a 120-min panel session per day for six samples. In the evaluation sessions, samples were served one at a time to each panelist. After individual evaluation of each sample, a discussion was initiated by the panel leader, during which consensus scores for all attributes were reached. Consensus method has been used in other recent studies (Di Donfrancesco et al., 2012; Sanchez & Chambers, 2015; Sharma et al., 2020; Vázquez-Araújo et al., 2014). The consensus approach made it possible to add new attributes during sample evaluation: if a new attribute was identified, the panel would discuss about it, and added to the list if all agreed with it (Chambers, 2018). The final list of aroma attributes, definitions and references were shown in Appendix B. In all sessions, steamed towels were provided for panelists to cleanse their olfactory pathways.

Appearance, Texture and Skin Feel Evaluation

The testing attributes, references, definitions and protocols of sample preparation, skin preconditioning, preparation of test sites, sample delivery and application were adopted from the ASTM Standard Guide for Descriptive Analysis of Skin Creams and Lotions -technical assessor approach with slight modifications (ASTM Standard E1490-19, 2019). Before the test, a 1-min wash with unscented mild soap and 5-min dry-out period were carried out on panelists' forearms and hands. Test sites were marked as 44-mm circles on the inner forearms of each assessor using the bottom edge of a 4 oz paper cup. Each sample was applied and tested in a different test site on panelists' forearms. Prior to sample application, the temperature of each test site was recorded for each panelist using a portable infrared thermometer (Thermo Fisher Scientific, Waltham, MA). In-between samples, unscented wipes and paper towels were provided for panelists to clean their forearms and fingers. The temperature of the testing room was controlled at 21 ± 1 °C, and relative humidity was maintained at about 30%.

Prior to sample evaluation, two 90-min orientation sessions were completed by all panelists during which testing protocols, attributes, definitions, and references were reviewed using 8 of the 12 testing samples. After the orientation session, samples were evaluated in three

120-min sessions (days) with four samples being evaluated per session (day). Appearance, texture and skinfeel (pick-up, rub-out, afterfeel) were evaluated based on ASTM guide (ASTM Standard E1490-19, 2019). A scale ranging from 0-15 with 0.5 increments was used for panelists to rate the intensity of each attribute using a consensus approach, except for absorbency. Absorbency was defined as the number of additional rubs (after 20 rubs) at which the product loses wet and a resistance to continue is perceived (ASTM Standard E1490-19, 2019). As a result, the number of additional rubs was recorded for individual panelists. Specific instructions were given regarding sample application and manipulation in each stage of evaluation adapted from the ASTM guide (ASTM Standard E1490-19, 2019).

Consumer Home Use Test

Samples

Seven hand cream samples that covered a wide range of aroma and textural profiles (four quadrants and center of PCA biplots) were selected for a consumer home use test (Figure 3.1). The sample selection was based on the sensory map created from the descriptive analysis of twelve hand cream samples. Hand cream samples were portioned into 5 ml cosmetic sample jars with film sealers and lids. All samples (in the jars) were labeled with random three-digit blinding codes.

Figure 3.1. Seven hand cream samples used in the consumer home-use study

Participant Recruitment

One hundred female participants between the ages of 18 and 54 were recruited from the Kansas City area via the consumer database of Sensory and Consumer Research Center at K-State Olathe. Participants had to be regular users of hand cream products (used hand creams at least once a day currently), have no skin allergies, and be willing to try different hand creams. A mix of skin types were recruited (22% dry, 62% normal-dry, 3% combination, 13% normal, 0% oily). The research was approved by the Institutional Review Board for Protection of Human Subject of Kansas State University (IRB # 10062).

Data Collection

The study was conducted during mid-March (Spring) of 2021, with the average daily temperature ranging between 4 °C to 15 °C in Kansas City. The average relative humidity during the testing week was 66.7%. The seven hand cream samples were tested in a home-use setting during a consecutive 7-day period with one sample to be evaluated per day. The order of sample evaluation followed a balanced rotation in a Latin Square design. On each testing day, participants were asked to use a different sample as many times as they normally would, and then complete an online questionnaire by the end of the day regarding that sample. Participants were required to suspend the use of their own hand creams throughout the duration of the study. Data collection was completed using Compusense Cloud Software (Compusense, Inc., Guelph, Ontario, Canada).

Questionnaire

Even though participants used each sample during the day, they were asked to have the sample in front them while completing the questionnaire so that they could retry as needed. The questionnaires included questions relevant to acceptance, emotions, attribute evaluation for each sample, and non-sample related questions such as consumer behavior towards the product category and demographics. The questionnaire took between10 and 15 minutes to complete. In the questionnaires, participants first rated their overall liking for the sample on a 9-point hedonic scale. Then, they were asked to rate their emotional responses on a 5-point scale (1 = not at all; 2 = slightly; 3 = moderately; 4= very much; 5 = extremely) using the emotion lexicon generated for skincare products in Chapter 2 (Table 3.2). This emotion lexicon contains a total of 42 terms including 27 positive terms and 15 negative terms. In addition, aroma, texture, afterfeel liking, as well as the intensity of aroma (JAR scale), thickness (JAR scale) and greasy (5-point intensity scale) were asked. One CATA question was included in the questionnaire to get an idea of description of these samples from a consumer perspective. The CATA terms were selected from previous studies regarding creams and lotions and online reviews regarding hand creams (Parente et al., 2011). In the non-sample related section of the questionnaire, participants were asked to rate twelve statements regarding their attitudes and behaviors on a 7-pt Likert scale.

Table 3.2. The emotion lexicon used for the hand cream HUT

Positive Emotions	Negative Emotions
Awake (+)	Bad/Unpleasant (-)
Balanced (+)	Bland (-)
Calm/Soothed (+)	Disappointed (-)
Clean (+)	Disgusted (-)
Comfortable (+)	Duped/Deceived (-)
Confident (+)	Flawed (-)
Fresh (+)	Incomplete (-)
Happy (+)	Insecure/Unconfident (-)
Healthy (+)	Self-conscious/Uncomfortable (-)
Luxurious (+)	Sloppy/Messy/Unkempt (-)
Natural (+)	Unhealthy (-)
Neat (+)	Unimpressed (-)
Youthful (+)	Unnatural (-)
Nourished (+)	Dissatisfied (-)
Pampered (+)	Frustrated (-)

Perfect/Flawless (+)	Regret (-)
Pleasant (+)	
Pleased/Satisfied (+)	
Pretty/Beautiful/Attractive (+)	
Put-together/Polished (+)	
Ready (+)	
Refreshed/Rejuvenated/Energized (+)	
Relaxed (+)	
Relieved (+)	
Vibrant (+)	
Annoyed/Irritated (-)	

Data Analysis

Mapping the Sensory Space of Hand Creams

Consensus-based sensory data from descriptive analysis was analyzed with principal component analysis (PCA) using correlation matrix. Aroma profiling map and map for appearance & texture & skinfeel were constructed separately.

Segmenting Hand Cream Consumers Based on Liking Patterns

Agglomerative Hierarchical Clustering (AHC) using Euclidian distance and Ward's method was performed on standardized consumer overall liking scores (centered and reduced across participants to reduce scale usage effect) for all seven hand cream samples. The dendrogram from AHC suggested a three consumer-cluster solution. To validate if the three segments had different rating patterns across products, two-way analysis of variances (ANOVA) were conducted on overall liking, product attribute liking, attribute intensity ratings and emotion ratings, considering main effects as sample, segment, and their interaction effect. The interaction effects will help us understand if rating patterns for the seven samples were dependent of consumer segments. In addition, between-segment differences in non-sample related consumer behavior were assessed by performing one-way ANOVA on the statement-agreement data and Chi-square test of independence on self-reported skin type data.

Identifying Sensory Drivers of Liking

Aroma drivers of liking, texture drivers of liking and afterfeel drivers of liking were explored separately. External preference mapping was performed on a cluster level by modeling the three consumer segments' aroma/ texture/ afterfeel liking scores with the samples' coordinates of the first and second dimensions from PCA on aroma or texture data. To predict the attribute liking of each consumer cluster, linear, circular, elliptical, and quadratic models were tested at a significant level of 0.1. This analysis was carried out using the PREFMAP function in XLSTAT (Addinsoft, New York, NY).

Exploring Emotion, Sensory and Consumer Perception Associations

To explore likings and emotions within each segment, overall liking and emotion ratings from each consumer segment were independently analyzed with ANOVA mixed model considering sample as fixed effect and participants as random effect, and Tukey's HSD post hoc test was performed on any significant sample effects detected by ANOVA. PCA was carried out on the mean ratings of emotions that significantly differentiated samples. As supplement variables, sensory data and overall liking were projected into the product-emotion space. For each consumer segment, Multiple Factor Analysis (MFA) was performed on attribute intensities from the descriptive analysis and consumer CATA frequency data to identify the relationship between sensory characteristics of the samples and consumer perception, with overall liking as a supplement variable.

All data analysis was conducted using XLSTAT (Addinsoft, New York, NY). A significance level of 0.05 was considered in this study, except for the model selection procedure in external preference mapping which used a significant level of 0.1.

Results and Discussion

Sensory Space of Hand Creams

Table 3.3 presents the consensus-based intensity scores by descriptive analysis on aroma, appearance, and texture attributes for the twelve hand cream samples. The results suggested that these samples differed widely in their sensory characteristics. The sensory spaces based on aroma attributes and appearance and texture & skinfeel attributes were constructed by PCA.

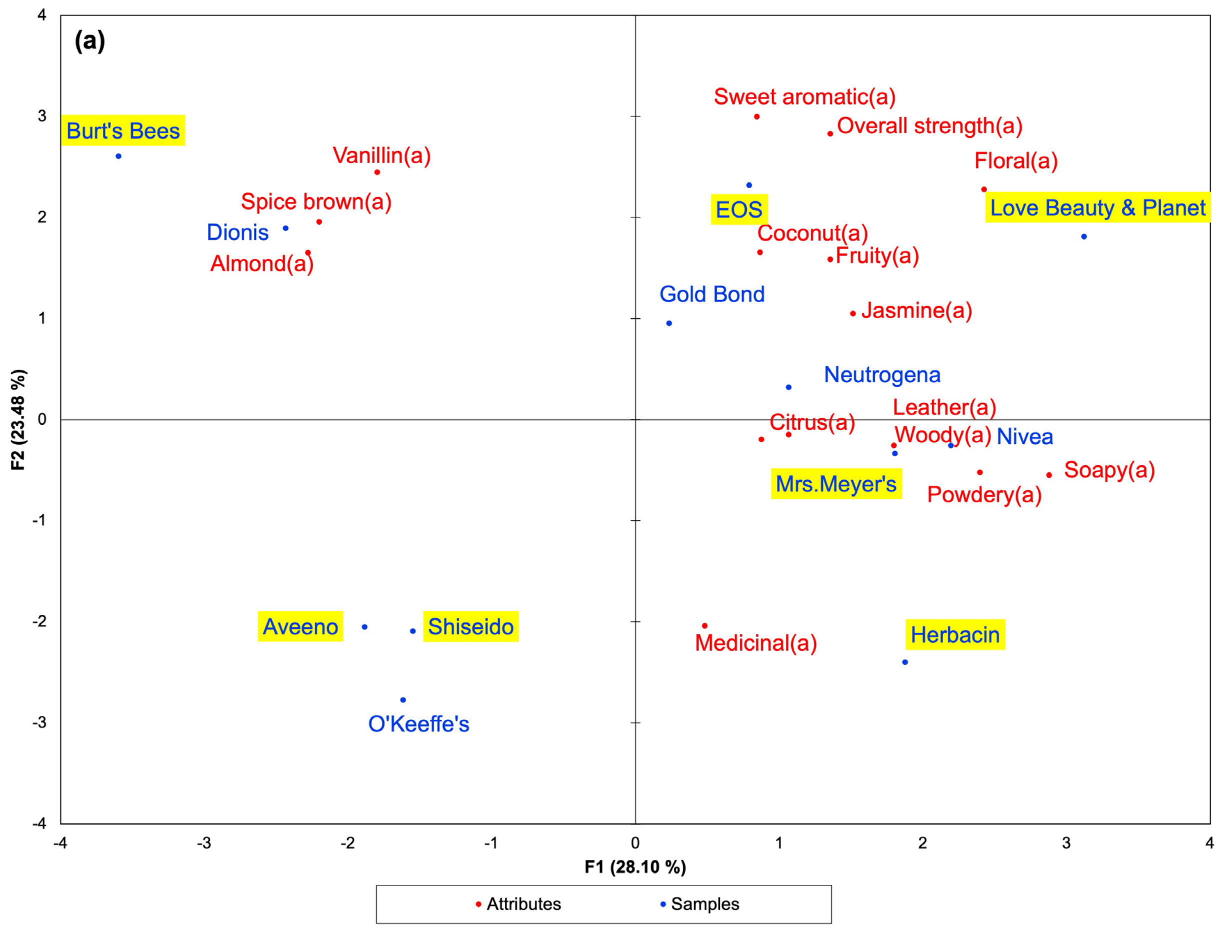
Overall, the twelve hand cream samples had very discrete aroma characteristics. A total of 15 aroma attributes were identified. The only attributes used to describe all samples were overall strength and sweet aromatic. Intensities of overall aroma strength ranged from very weak (Aveeno, Shiseido and O'Keeffe's, these were labeled as fragrance-free samples) to quite strong (Nivea, Love Beauty & Planet, Burt's Bees). Sweet aromatic varied from quite low (Aveeno, Shiseido and O'Keeffe's) to relatively higher (Gold Bond, EOS and Love Beauty & Planet). For those samples that were not fragrance-free, each sample can be characterized by its key aroma notes: Burt's Bees- almond and spice brown, Dionis-vanillin and spice brown, EOS- coconut, fruity and sweet aromatic, Gold Bond- fruity and sweet aromatic, Herbacin- powdery and medicinal, Love Beauty & Planet- floral and jasmine, Mrs. Meyers- soapy and citrus, Nivea- soapy and leather, Neutrogena-floral, fruity, and powdery. Based on the first two dimensions of PCA (Figure 3.2 (a)), these samples can be seen in five groups for their aroma profiles: the low-aroma group (Aveeno, Shiseido and O'Keeffe's), the mild-powdery & medicinal group (Herbacin), the spice-aroma group (Burt's Bees and Dionis), the fruity & floral group (Gold

Bond, EOS and Love Beauty & Planet) and the powdery & soapy group (Nivea, Mrs. Meyer's and Neutrogena).

For appearance, texture and afterfeel evaluation, several attributes such as gloss, stickiness, oil, grease, and wax were identified at multiple points throughout product application. The residue each sample left on skin after rub-out was characterized as either oil (baby oil), grease (lanolin) or wax (beeswax). The first two dimensions of PCA (Figure 3.2 (b)) on the twenty-four attributes explained 73.3% of the variance among samples. The first principal component (PC) explaining 41.54% of the total variance represented characteristics related to viscosity of samples. This PC was negatively correlated with spreadability (rb), slipperiness (af immediate), wetness (rb) and positively correlated with thickness (rb), firmness (p), integrity of shape (ap). The second PC explaining 31.76% of the variance represented the absorbency and residue characteristics of samples. It was negatively associated with absorbency (rb), wax (rb), wax (af immediate and after 5 min), and positively associated with grease (rb, af immediate and after 5 min), gloss (af immediate and after 5 min). The sensory characteristics and skin performances of skincare products are the results of their ingredients such as different emollients and the way they were formulated (Boinbaser et al., 2015; Parente et al., 2008; Wortel & Wiechers, 2000). Of these twelve samples, Nivea and Burt's were characterized as high thickness, and high grease during and after rub-out. O' Keeffe's and Love Beauty & Planet also had high thickness and created high waxy skin feel. Herbacin, EOS and Gold Bond were on the waxy side but not as thick. Mrs. Meyer's was the most runny, wet, easy to spread of all and had relatively oily texture and residue, explaining the reason why it left skin most slippery after rubbing. Dionis, Neutrogena and Shiseido were also on the less thick side of the map, but

Shiseido and Neutrogena were characterized as slightly grease and high cohesiveness. Aveeno was the sample that had intermediate levels of all appearance and texture attributes.

Based on the sensory maps generated from descriptive analysis **(Figure 3.2)**, Burt's Bees, Love Beauty & Planet, Herbacin, EOS, Aveeno, Shiseido, and Mrs. Meyers were selected for consumer home use test. The sensory results of these seven samples were used to identify sensory drivers of liking and emotional associations.



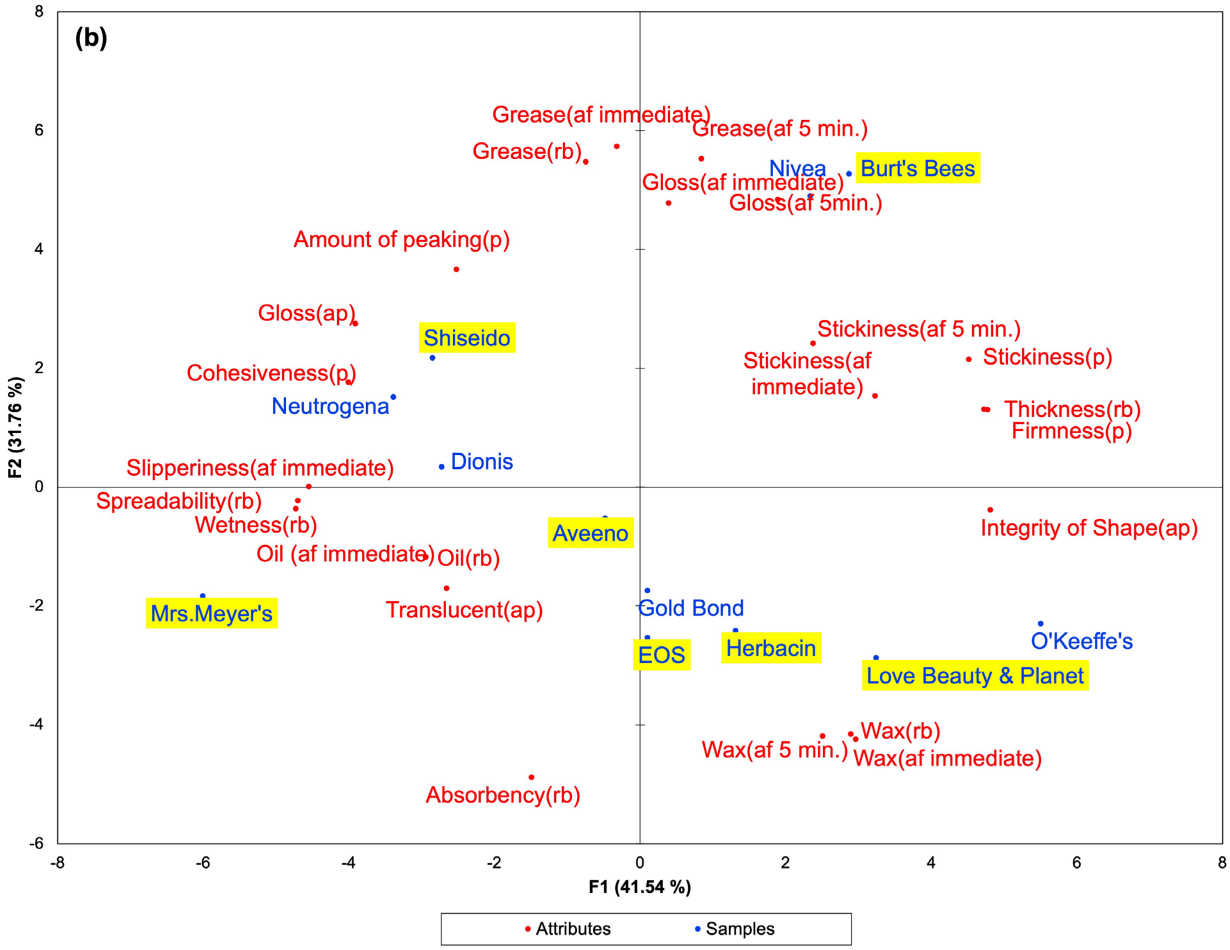


Figure 3.2. Perceptual maps of hand creams represented by the first two dimensions of PCA on aroma attributes (a) and appearance, texture and afterfeel attributes (b). a = aroma, ap = appearance, p = pick-up, r = rub-out, af = afterfeel, af 5 min.= afterfeel after 5 min. Samples highlighted were selected for consumer HUT

Table 3.3. Attribute intensity scores for the descriptive analysis for hand cream samples (0–15-point scale with 0.5 increments except for absorbency (rb)). a = aroma, ap = appearance, p = pick-up, r = rub-out, af = afterfeel, af 5 min.= afterfeel after 5 min

Attributes	Aveeno	Burt's Bees	Dionis	EOS	Gold Bond	Herbacin	Love Beauty & Planet	Mrs.Meyer's	Neutrogena	Nivea	O'Keeffe's	Shiseido
Overall strength (a)	2.5	9	8	7	7	5	9.5	8	8	10	2	2
Soapy (a)	0	0	0	1.5	1.5	4	3	5	3	5	1	1
Floral (a)	0	2.5	3	5	3	3	6	4	5	5	0	1
Jasmine (a)	0	0	0	0	0	0	4	0	0	0	0	0
Fruity (a)	0	0	0	4.5	5.5	0	2	3	3	0	0	0
Citrus (a)	0	0	0	0	0	0	0	4	0	0	0	0
Almond (a)	1	5.5	2	0	0	0	0	0	0	0	0	0
Coconut (a)	0	0	0	3	0	0	1	0	0	0	0	0
Spice brown (a)	0	4	3	0	0	0	0	0	0	0	0	0
Sweet aromatic (a)	2	4	4	5	5.5	3	5	3	3	3	2	2
Vanillin (a)	0	2.5	3	1.5	0	0	0	0	1	0	0	0
Medicinal (a)	0	0	0	0	0	2	0	0	0	0	1	0
Powdery (a)	1	0	1	1	0	4	2	2	3	3	0	0
Woody (a)	0	0	0	0	0	3	3	0	0	0	0	0
Leather (a)	0	0	0	0	0	0	0	0	0	3	0	0
Integrity of Shape (ap)	12	12	11	11	11.5	12	12.5	10	11	12	13	10.5
Gloss (ap)	10	10	13	8.5	9	9	3	11	12	10	0	11
Translucent (ap)	2.5	0	0	0	3	3	0	3	1.5	0	0	2
Firmness (p)	6	9	4.5	5.5	6	7	7	4	4.5	8.5	10	5
Stickiness (p)	6.5	9.5	5.5	6	6	7	7.5	4.5	5	8.5	8	5
Cohesiveness (p)	4	5	6.5	6	3	4	4	6.5	6	4.5	2	7
Amount of peaking (p)	3.5	5.5	4.5	4.5	4	4	3	5	4.5	5.5	2	8

Attributes	Aveeno	Burt's Bees	Dionis	EOS	Gold Bond	Herbacin	Love Beauty & Planet	Mrs.Meyer's	Neutrog ena	Nivea	O'Keeffe's	Shiseido
Wetness (rb)	4.5	4	5.5	5	4.5	4.5	3.5	7.5	6.5	4	4	5.5
Spreadability (rb)	7.5	6	8	7	7.5	6.5	4.5	9.5	8.5	5.5	6	8.5
Thickness (rb)	5	9	5	5.5	6	6.5	8.5	3	4	7.5	8	5.5
Wax (rb)	0	0	0	4	3	2.5	3	0	0	0	4	0
Grease (rb)	2	3.5	2.5	0	0	0	0	0	3	3.5	0	3
Oil (rb)	0	0	0	0	0	0	0	3	0	0	0	0
Absorbency (rb)*	0.052	0.024	0.048	0.059	0.061	0.059	0.078	0.063	0.048	0.014	0.038	0.049
Gloss (af immediate)	0	3.5	1	0	2	0	0	0	3.5	3	1	1.5
Stickiness (af immediate)	3	3.5	2	2	3	3	2	2	2	3	4	3
Slipperiness (af immediate)	5.5	4	6	3.5	5.5	4	3	7	7	3	3	5
Wax (af immediate)	0	0	0	3	3	3	3	0	0	0	3.5	0
Grease (af immediate)	1	3	1.5	0	0	0	0	0	2	3	0	2.5
Oil (af immediate)	0	0	0	0	0	0	0	2	0	0	0	0
Gloss (af 5min.)	0	1.5	0	0	0	0	0	0	0	1.5	0	0
Stickiness (af 5 min.)	1	3	0	1	1	1	1	1.5	1	1.5	2	1
Wax (af 5 min.)	1	0	0.5	2	0	2	1.5	0.5	0	0	2.5	0
Grease (af 5 min.)	0	2.5	0	0	0	0	0	0	1	3	0	1.5
Oil (af 5 min.)	0	0	0	0	0	0	0	0	0	0	0	0

**Absorbency during rub-out was evaluated individually. The values presented in the table for absorbency were 1 divided by the average number of rubs among the panelists at which the product lost wetness on skin. The larger the value of absorbency, the easier it was for the product to be absorbed by skin.*

Consumer Segmentation from Home Use Test

Three consumer clusters were identified by AHC performed on the standardized overall liking scores collected from the seven samples evaluated in the HUT (Figure 3.3). Table 3.4 provided the mean overall liking scores for each product within each consumer cluster. Of the seven samples tested, all three consumer clusters liked EOS (coconutty, waxy and intermediate thickness) and disliked Burt's Bees (strong almond, thick and greasy). For the other five samples, the three clusters had different overall preference patterns. Cluster 1 preferred Love Beauty & Planet (floral, thicky, waxy) but disliked Mrs. Myers (citrus & soapy, thin, wet, and oily). However, cluster 2 liked the low-scented Herbacin and disliked the high-scented Love Beauty & Planet. Cluster 3, in contrast to cluster 1 and 2, liked all samples except for the unscented samples-Aveeno and Shiseido. The different preference patterns between different clusters seemed to be affected by consumers' preferences for different aroma and texture experiences.

Two-way ANOVAs considering sample, cluster and interaction effect on overall liking, attribute liking, intensity scores, and emotion ratings (Appendix D& Appendix E) confirmed the appropriateness of the three-cluster solution for consumer segmentation. Cluster*sample interaction effects were significant for overall liking, attribute likings, aroma intensity rating except for thickness, and ratings of the 42 emotion responses. This indicated the three clusters of different product liking patterns also differed in attribute liking as well as emotional associations with different hand creams. However, they still had the same perception/understanding on the thickness attribute.

Usage and preference behavior differences among clusters were revealed partially by analyzing the statement agreement data (Figure 3.4). The three clusters differed significantly on their attitudes towards four statements describing their preferences for aroma and texture of hand creams. Consumers in cluster 1 were more likely to prefer scented hand creams than non-scented, and thick texture than thin texture. Consumers in cluster 3 liked their hand creams to be scented and disliked scent-free ones. There was no strong indication on preference for thin or thick texture for cluster 3. Consumers in cluster 2 were more likely to like unscented samples than cluster 1 and 3. Despite their differences for sensory preferences, no significant associations were identified between clusters and skin type as well as clusters and consumer demographics (Appendix C).

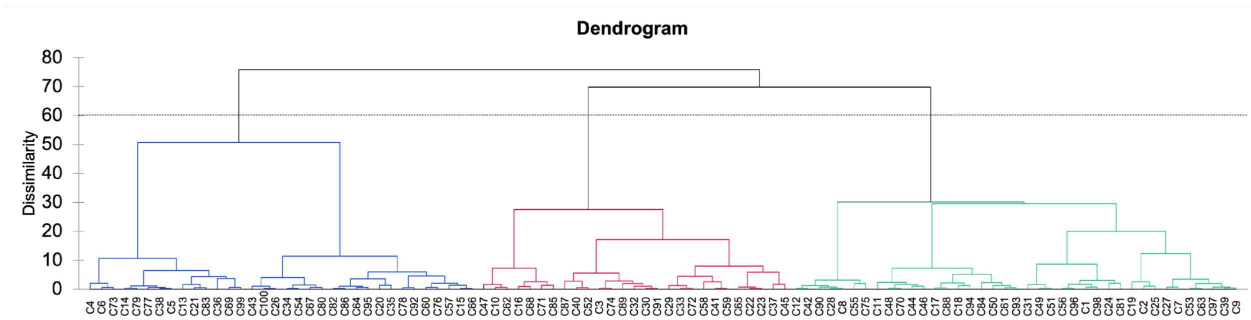


Figure 3.3. AHC dendrogram grouping consumers into three clusters based on standardized overall liking scores

Table 3.4. Mean overall liking scores of the seven hand cream samples for each consumer cluster (9-point hedonic scale, 1= dislike extremely, 5= neutral, 9= like extremely)

Cluster	Aveeno	Burt's Bees	EOS	Herbacin	Love Beauty & Planet	Mrs. Meyer's	Shiseido
1 (N=39)	6.8 ab	5.5 c	7.5 a	5.3 cd	7.4 a	4.2 d	6.1 bc
2 (N=27)	6.9 ab	4.0 c	7.0 ab	7.3 a	3.9 c	5.7 b	6.4 ab
3 (N=34)	4.0 b	5.1 b	7.3 a	6.9 a	7.5 a	7.2 a	4.9 b

Means with different letters were significantly different within each consumer cluster (p<0.05). Within each cluster, most liked samples (green) and least liked samples (yellow) were highlighted.

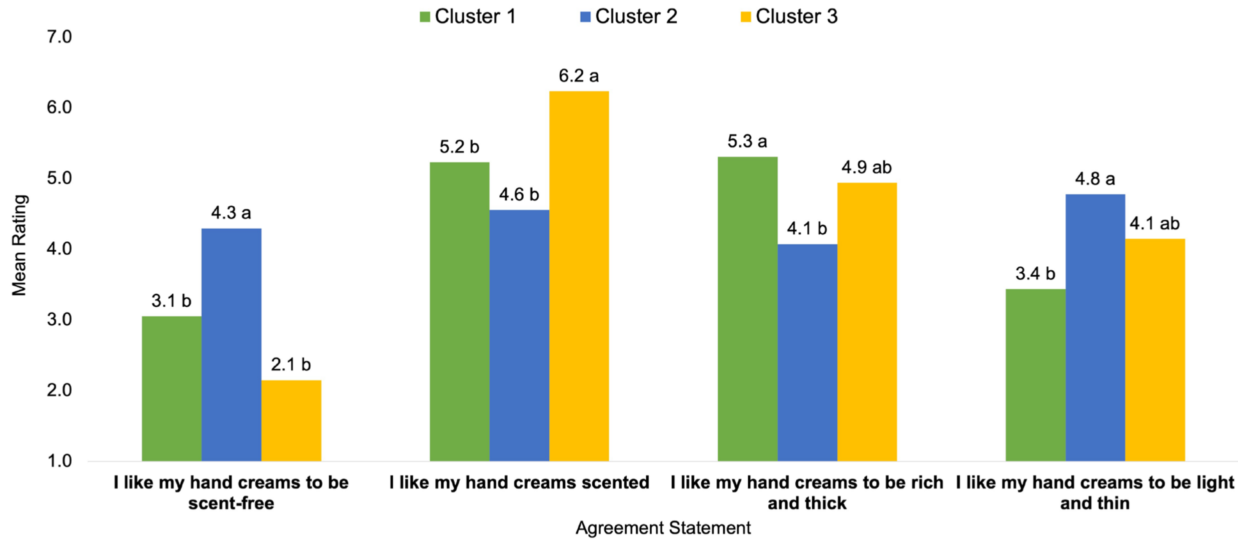


Figure 3.4. Non-sample related agreement statements that significantly differentiated the three clusters (7-pt Likert scale, 1=strongly disagree, 7= strongly agree). Means with the same letter designations were not statistically different within each statement (alpha=0.05)

Sensory Drivers of Liking for Hand Creams

Instead of performing preference mapping based on overall liking, sensory drivers of liking for hand creams were explored based on aroma liking, texture liking and afterfeel liking separately. The seven hand cream samples had very discrete aroma profiles-they seldom shared the same aroma attributes. Combining discrete aroma profiles together with a diverse range of texture profiles would create a sensory configuration that was hard to be explained by overall liking. As a result, external preference maps (Figure 3.6) were obtained from regression of mean aroma liking scores of the three clusters on the first two dimensions of the PCA on aroma attributes, and regressions of mean texture and afterfeel liking scores of the three clusters on the appearance & texture & skinfeel space. Drivers of appearance liking was not considered in this study as appearance liking did not seem to be widely different between samples and consumer clusters (Figure 3.5).

The first two dimensions of the sensory aroma map represented 62.79% of the total variance within samples. External preference mapping on the first two dimensions for aroma (Figure 3.6 (a)) fitted cluster 1 with a quadratic model, cluster 2 with an elliptical model, and cluster 3 with a vector model. The quadratic model for cluster 1 provided the anti-ideal point located close to the center of the map and sample Mr. Meyer's. This made the relationship complex and hard to identify the specific attributes this cluster preferred. But according to Figure 3.5, this group gave Mrs. Myers and Herbacin lower aroma liking scores than other samples, suggesting they were very likely to dislike citrus, medicinal, soapy, and powdery aroma. However, it seems that this group would still prefer something with scent rather than non-scented products. The elliptical model fitted for cluster 2 obtained a saddle point near the center of the map, representing a threshold where aroma acceptance was decreasing on the direction to sample Love Beauty & Planet and Mrs. Meyer's, but was increasing on the direction of the rest of samples. Combing the aroma liking patterns across the seven samples, most consumers in this cluster appeared to prefer non-scented or low-scented samples, and seemed to dislike strong aroma intensities, especially if the strong aromas were floral or powdery & soapy. The vector model for cluster 3 suggested the direction of aroma liking appearing to be towards the positive side of both PCs. This cluster's aroma liking was driven by strong, fruity, floral, and sweet aroma attributes.

The first two dimensions of the sensory texture map represented 79.29% of the total variance within the samples. External preference mapping on texture liking and afterfeel liking provided similar results due to the high correlation between the likings of these two attributes (Figure 3.6 (b) and (c)). Vector models were obtained for the prediction of both texture and afterfeel likings of cluster 1 and cluster 3. The direction of texture and afterfeel liking for both clusters was towards the positive side of PC 2, indicating the drivers were high waxy afterfeel and high absorbency. The differences between these two clusters were cluster 1 preferred higher intensities of thickness, firmness, and integrity of shape, while cluster 3 preferred a less thick and firm texture. Cluster 2 was fitted with circular models with ideal points identified on the map for both texture and afterfeel liking. The texture profiles of EOS, Aveeno and Herbacin were all considered as ideal by this group. Compared to cluster 1 and 3, cluster 2 liked a less thick, less waxy hand cream but with a slightly more slippery afterfeel, a more wet and easier to spread texture.

Overall, the three clusters had different drivers of liking, while the afterfeel of high greasy, gloss and stickiness were found to be disliked by all clusters. Similar results were presented in the study of antiaging cream (Parente et al., 2011). Cluster 1 was defined as the thick & waxy texture liking group. The consumers in this group were predicted to like thick and waxy texture, waxy afterfeel and high absorbency for their hand creams. Aroma was less dominant compared to texture in driving their likings even though they did have specific aroma notes they disliked (citrus, medicinal and powdery). Cluster 2 was the mild scent & low-medium thickness liking group, who disliked strong aroma and extreme thick texture for their hand creams. Cluster 3, on the other hand, was the strong scent liking group. This group liked samples with strong intensities of scent, no matter if they were fruity, floral, sweet, spice brown, coconut, and they disliked the scent-free products. Both thin and thick texture (if not greasy) seemed to be acceptable for this group.

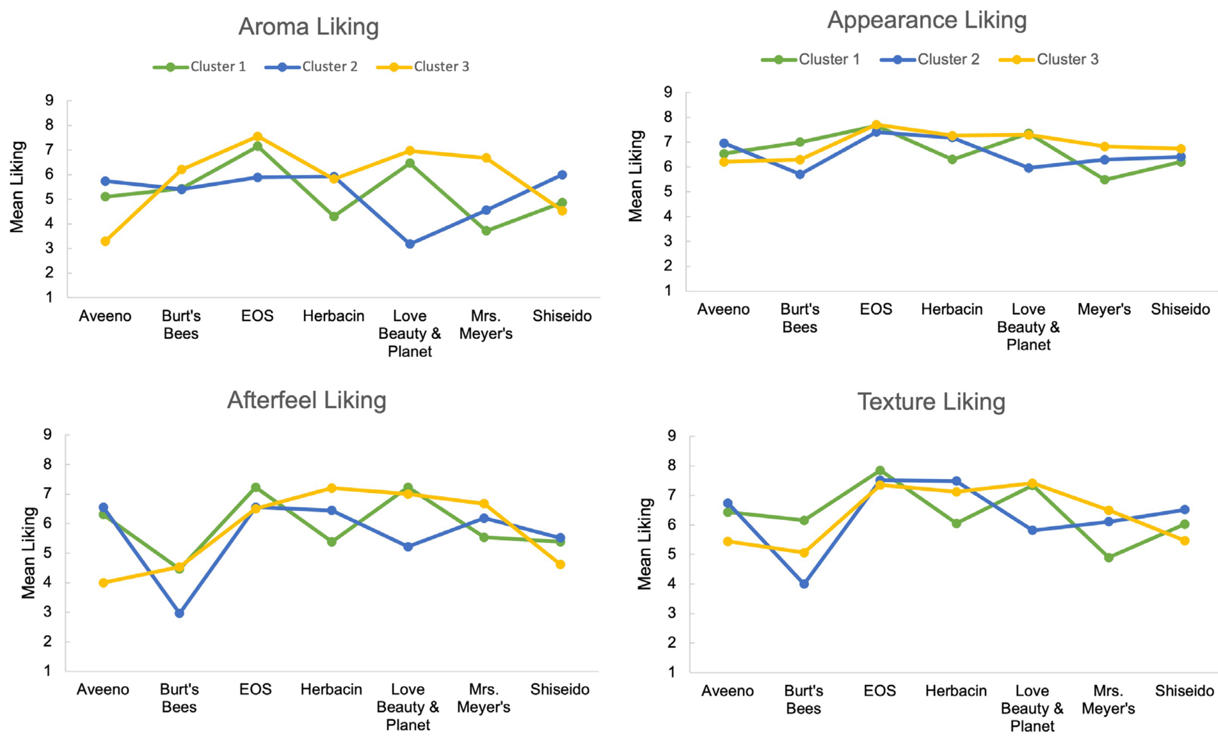
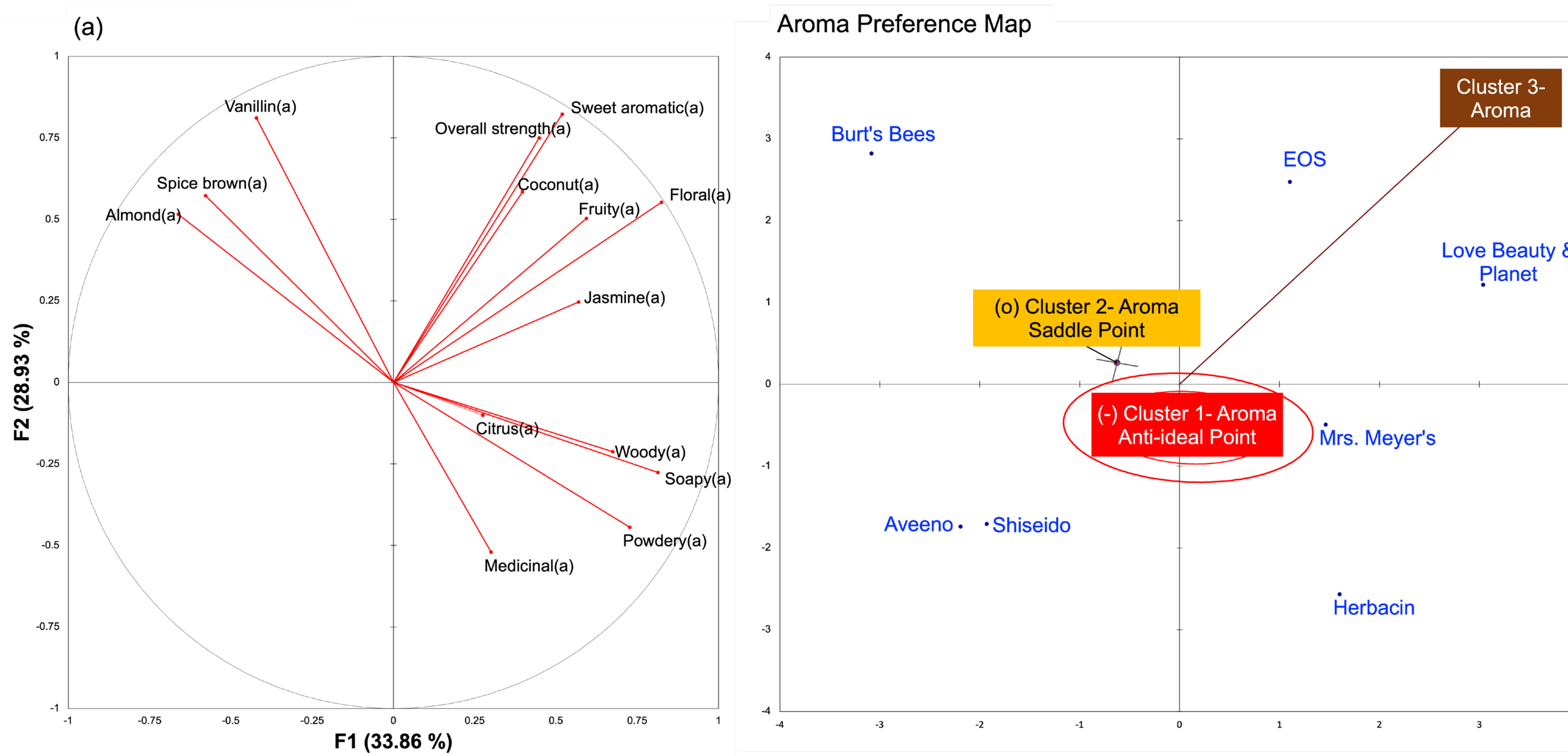
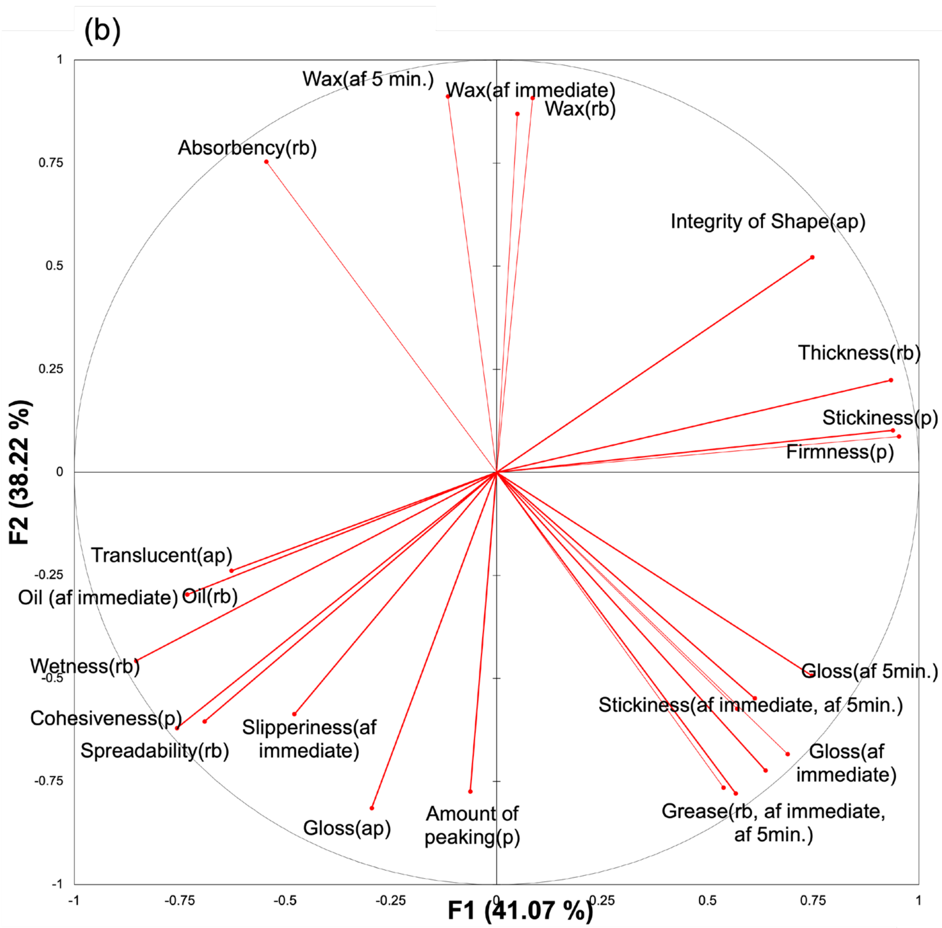
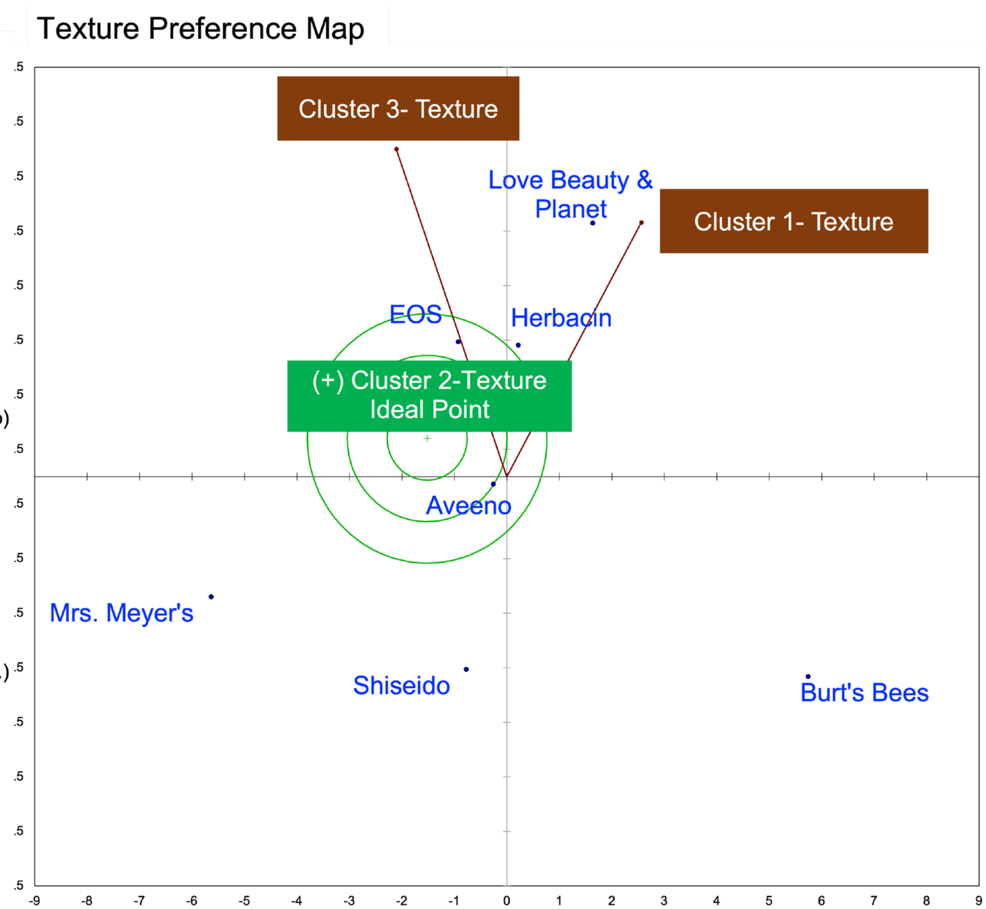


Figure 3.5. Mean attribute liking scores of the seven hand cream samples for each consumer cluster (9-point hedonic scale, 1= dislike extremely, 5= neutral, 9= like extremely)


(a)
Vanillin(a)
Sweet aromatic(a)
Overall strength(a)
Spice brown(a)
Coconut(a)
Floral(a)
Almond(a)
Fruity(a)
Jasmine(a)
Citrus(a)
Woody(a)
Soapy(a)
Powdery(a)
Medicinal(a)
F1 (33.86 %)
F2 (28.93 %)
Aroma Preference Map
Burt's Bees
EOS
Cluster 3- Aroma
Love Beauty &
Planet
(o) Cluster 2- Aroma Saddle Point
(-) Cluster 1- Aroma Anti-ideal Point
Mrs. Meyer's
Aveeno
Shiseido
Herbacin



(b)
Wax(af 5 min.)
Wax(af immediate)
Wax(rb)
Absorbency(rb)
Integrity of Shape(ap)
Thickness(rb)
Stickiness(p)
Firmness(p)
Translucent(ap)
Oil (af immediate)
Oil(rb)
Wetness(rb)
Cohesiveness(p)
Spreadability(rb)
Slipperiness(af immediate)
Gloss(ap)
Amount of peaking(p)
Gloss(af 5min.)
Stickiness(af immediate, af 5min.)
Gloss(af immediate)
Grease(rb, af immediate, af 5min.)
F2 (38.22 %)
F1 (41.07 %)



Texture Preference Map
Cluster 3- Texture
Cluster 1- Texture
Love Beauty & Planet
EOS
Herbacin
(+) Cluster 2-Texture Ideal Point
Aveeno
Mrs. Meyer's
Shiseido
Burt's Bees

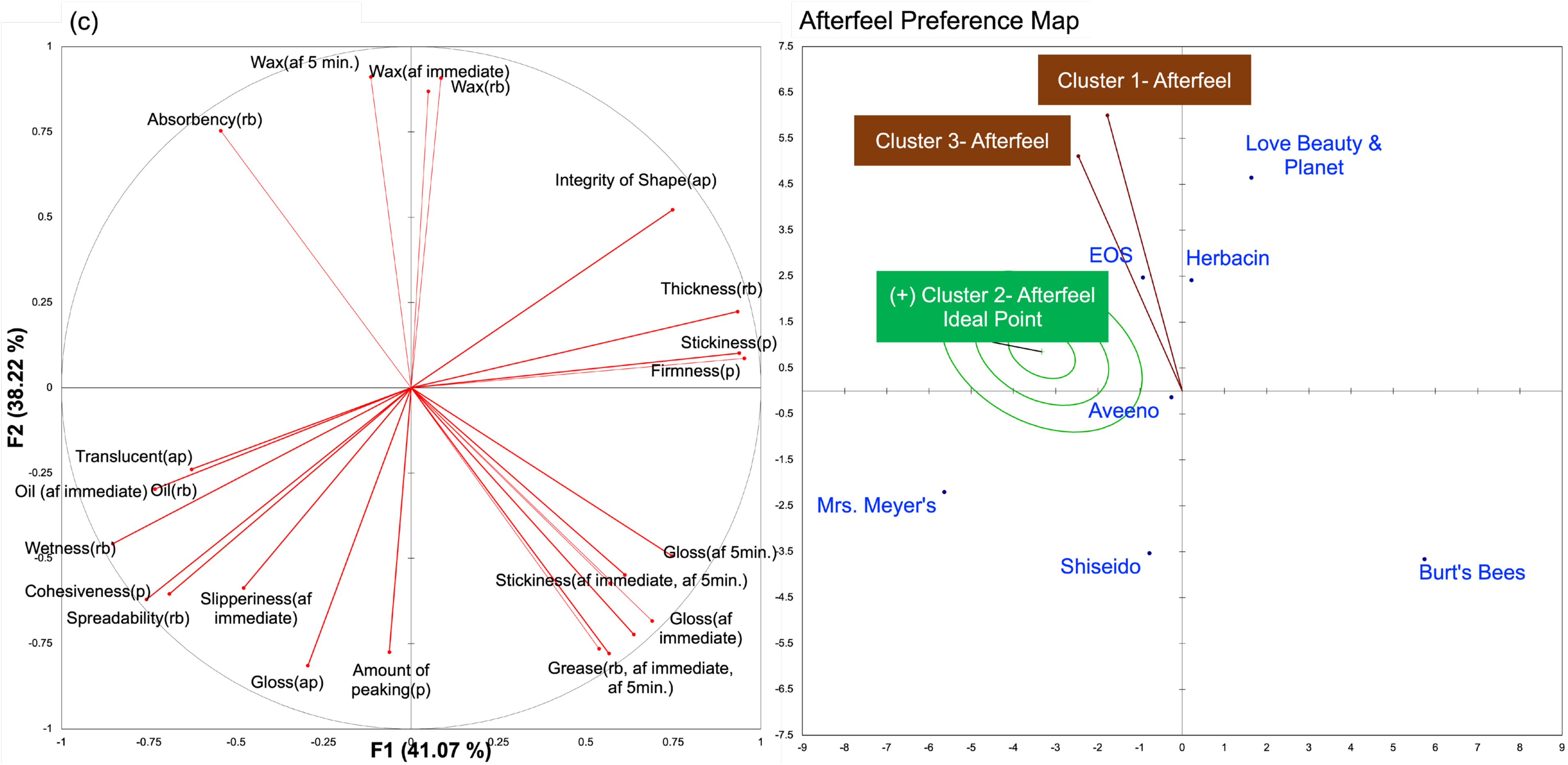


Figure 3.6. External preference mapping to identify sensory drivers of aroma (a), texture (b) and afterfeel (c) liking for the three consumer clusters

Consumer Perception, Emotion and Sensory Association

Emotion and Sensory Association

Analysis of variance conducted within each of the three clusters showed that most emotional terms were able to differentiate the different samples for each of the three clusters. The only few terms that did not show significant differences between samples were incomplete for cluster 1 and cluster 2, unhealthy for cluster 1 and cluster 3, and bland for cluster 2. Only significant emotional terms for each cluster were considered in studying emotion and sensory association.

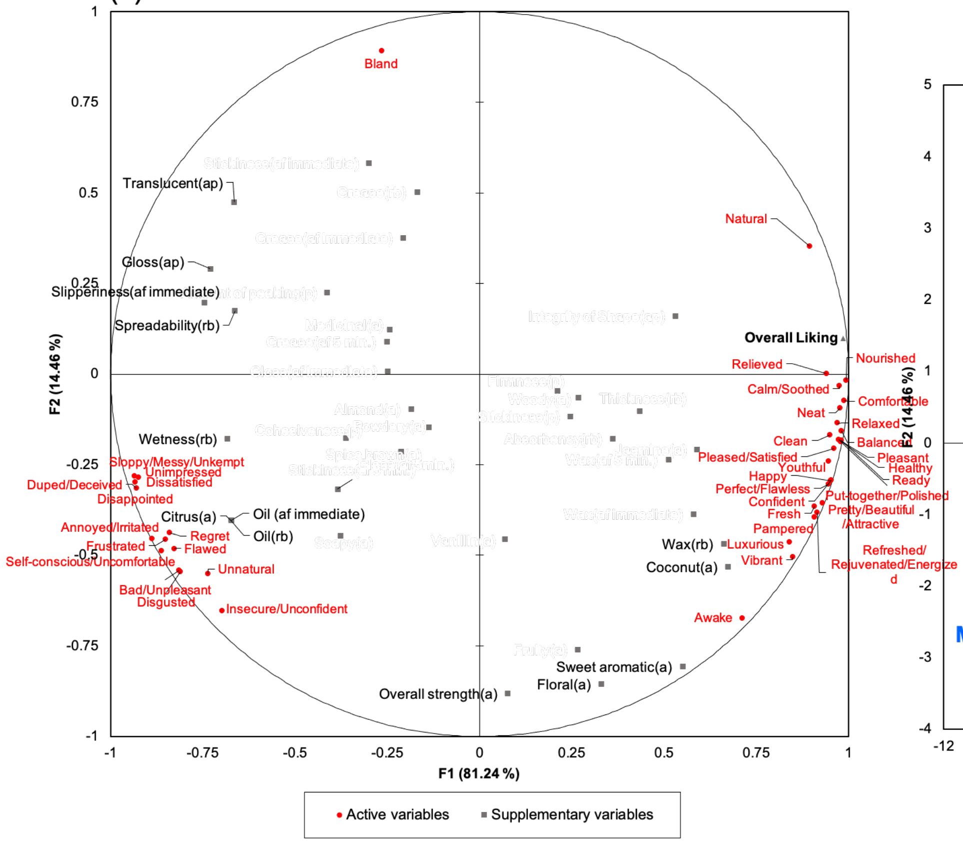
Emotional responses for the 7 hand creams representing different sensory profiles were mapped with PCA for each consumer cluster **(Figure 3.7)**. Sensory attributes and overall liking scores were used as supplementary data being projected into the product-emotion space based on correlation. For all three clusters, PC1s of the PCAs performed on emotional data all accounted for more than 80% of the data variance. This indicated that the tested hand cream products were mainly differentiated on the first dimension, represented by the valence of emotions (negative-positive). This dimension was also highly correlated with product overall liking. While the second dimensions accounted for only 4%-16% of the total variance of the data, they seemed to represent the arousal states of emotions (low-high). An overall pattern identified from the three configurations was the first dimension of emotion separated different products according to their overall liking within each consumer cluster, while the second dimension differentiated products based on their aroma characteristics. High intensities of certain aroma attributes seemed to elicit high-arousal emotions.

Looking at individual consumer clusters, it was observed that the desirable sensory attributes for each cluster elicited positive emotions, while undesirable attributes evoked negative emotions. The three clusters had different sensory drivers of likings, as a result, they associated sensory experiences with emotions in different ways. Only significant associations were considered in this exploration (Pearson correlation, $p < 0.05$). Cluster 1 was the waxy & thick texture liking group identified from preference mapping. For this group, waxy texture during rub-out was highly associated with several positive emotions including fresh, refreshed/rejuvenated/energized, and vibrant. While oil and wet texture evoked significant negative emotions such as annoyed/irritated, bad/unpleasant, disgusted, duped/deceived, flawed, self-conscious, frustrated and regret. In addition, attributes like translucent and gloss appearance, high spreadable texture, and slippery afterfeel were negatively correlated with several positive emotions. Aroma-wise, this cluster positively correlated coconut aroma with most positive feelings, except for natural; sweet aromatic with awake, luxurious, pretty/beautiful/attractive, refreshed/rejuvenated/energized, and vibrant; floral aroma with awake. Furthermore, lacking aroma strength would make consumers of this cluster feel bland (negatively correlated).

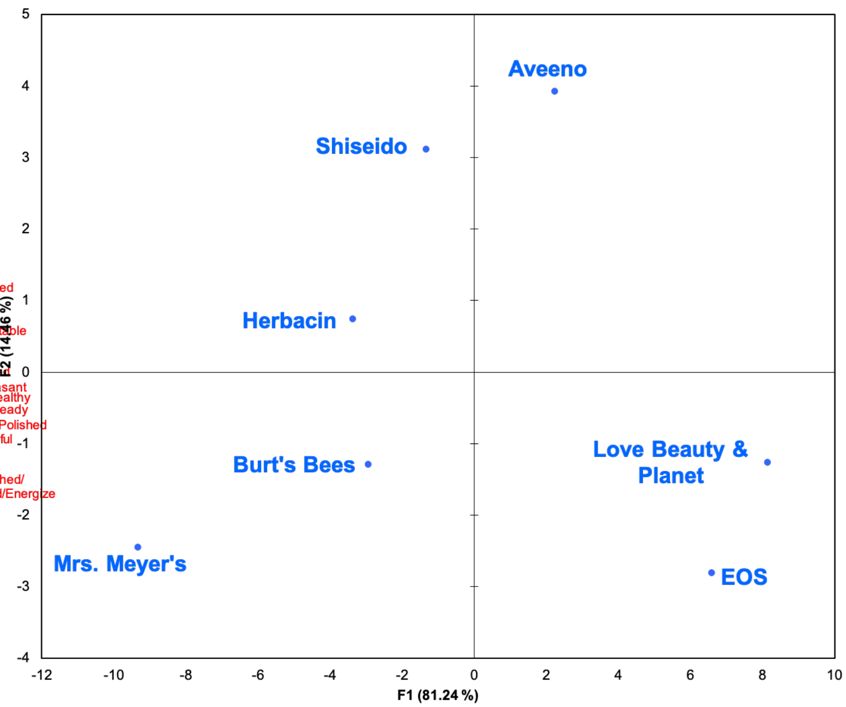
Cluster 2 was the group who preferred a texture of less thick and waxy but more spreadable, and mild aroma for their hand creams. For this group, hand creams of translucent appearance, low overall strength and low sweet aroma would make them feel natural. The strong jasmine perfumey aroma made this group feel unhealthy. In fact, for this group, aromas of high strength were negatively correlated with all positive emotions, and positively correlated with several high-arousal negative emotions such as bad/unpleasant, disgusted, and frustrated. Thick texture was negatively associated with the feelings of clean, fresh, pretty/beautiful and refreshed/rejuvenated/energized. In addition, this cluster related the texture of high stickiness, greasiness, and gloss afterfeel to the feelings of sloppy/messy/unkempt.

Cluster 3 was the aroma-driven group, who preferred high strength of scents in their hand creams. This group associated high overall aroma strength positively with awake, nourished, pretty/beautiful/attractive and vibrant; and negatively with bland. They also related floral aroma with almost every positive emotion measured. In addition, soapy aroma evoked the feelings of clean, relieved, comfortable, and natural. Fruity aroma made the group feel fresh, happy, and healthy. Sweet aromatic made them feel happy, healthy, luxurious, nourished, pampered and youthful. In terms of texture, this group related waxy texture with the feelings of luxurious, pampered, neat, put-together/polished, and wax afterfeel with feeling natural. Greasy texture during rub-out and greasy and sticky afterfeel evoked all kinds of negative emotions such as bad/unpleasant, annoyed/irritated, disgusted, flawed, insecure/unconfident, self-conscious, and sloppy/messy/unkempt.

(a) **Cluster 1**

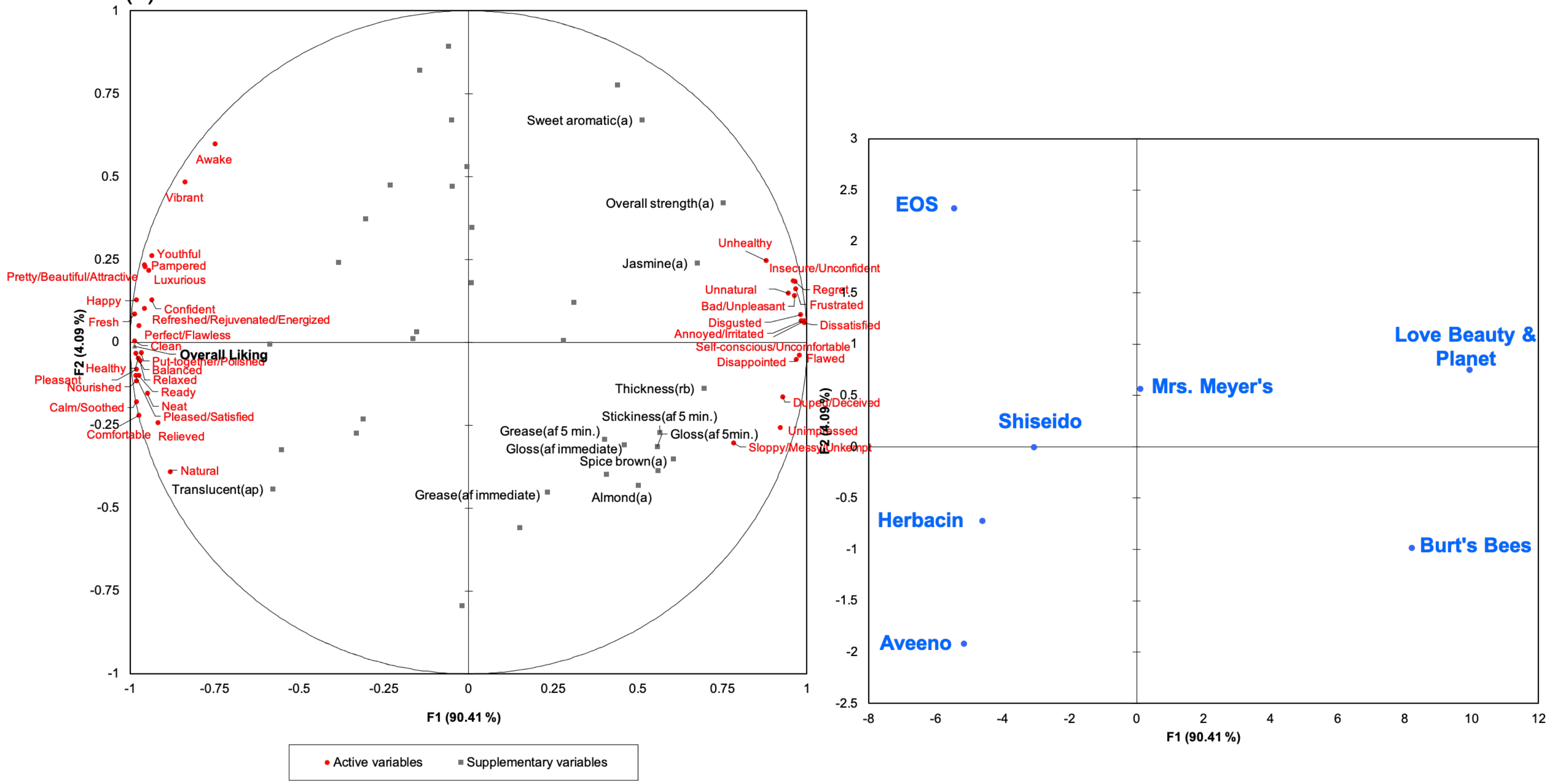


Bland
Natural
Translucent(ap)
Gloss(ap)
Slipperiness(af immediate)
Spreadability(rb)
Overall Liking
Relieved
Nourished
Calm/Soothed
Comfortable
Neat
Relaxed
Clean
Balanced
Pleased/Satisfied
Pleasant
Youthful
Healthy
Happy
Ready
Perfect/Flawless
Confident
Put-together/Polished
Fresh
Pretty/Beautiful /Attractive
Pampered
Wetness(rb)
Sloppy/Messy/Unkempt
Unimpressed
Duped/Deceived
Dissatisfied
Disappointed
Citrus(a)
Oil (af immediate)
Oil(rb)
Annoyed/Irritated
Regret
Frustrated
Flawed
Self-conscious/Uncomfortable
Unnatural
Bad/Unpleasant
Disgusted
Insecure/Unconfident
Wax(rb)
Luxurious
Vibrant
Coconut(a)
Refreshed/ Rejuvenated/Energized
Awake
Sweet aromatic(a)
Floral(a)
Overall strength(a)
F1 (81.24 %)
F2 (14.46 %)
1
0.75
0.5
0.25
0
-0.25
-0.5
-0.75
-1
0.25
0.5
0.75
Active variables
Supplementary variables



Aveeno
Shiseido
Herbacin
Burt's Bees
Love Beauty & Planet
Mrs. Meyer's
EOS
F1 (81.24 %)
F2 (14.46 %)
5
4
3
2
1
0
-1
-2
-3
-4
-12
-10
-8
-6
-4
2
4
6
8
10

(b) **Cluster 2**



Awake
Vibrant
Youthful
Pampered
Pretty/Beautiful/Attractive
Luxurious
Happy
Confident
Fresh
Refreshed/Rejuvenated/Energized
Perfect/Flawless
Clean
Overall Liking
Put-together/Polished
Healthy
Balanced
Pleasant
Relaxed
Nourished
Ready
Calm/Soothed
Neat
Pleased/Satisfied
Comfortable
Relieved
Natural
Translucent(ap)
Sweet aromatic(a)
Overall strength(a)
Jasmine(a)
Unhealthy
Insecure/Unconfident
Unnatural
Regret
Bad/Unpleasant
Frustrated
Disgusted
Dissatisfied
Annoyed/Irritated
Self-conscious/Uncomfortable
Disappointed
Flawed
Thickness(rb)
Duped/Deceived
Stickiness(af 5 min.)
Grease(af 5 min.)
Gloss(af 5min.)
Unimpressed
Gloss(af immediate)
Sloppy/Messy/Unkempt
Spice brown(a)
Grease(af immediate)
Almond(a)
F1 (90.41 %)
F2 (4.09 %)
Active variables
Supplementary variables
EOS
Love Beauty & Planet
Mrs. Meyer's
Shiseido
Herbacin
Burt's Bees
Aveeno

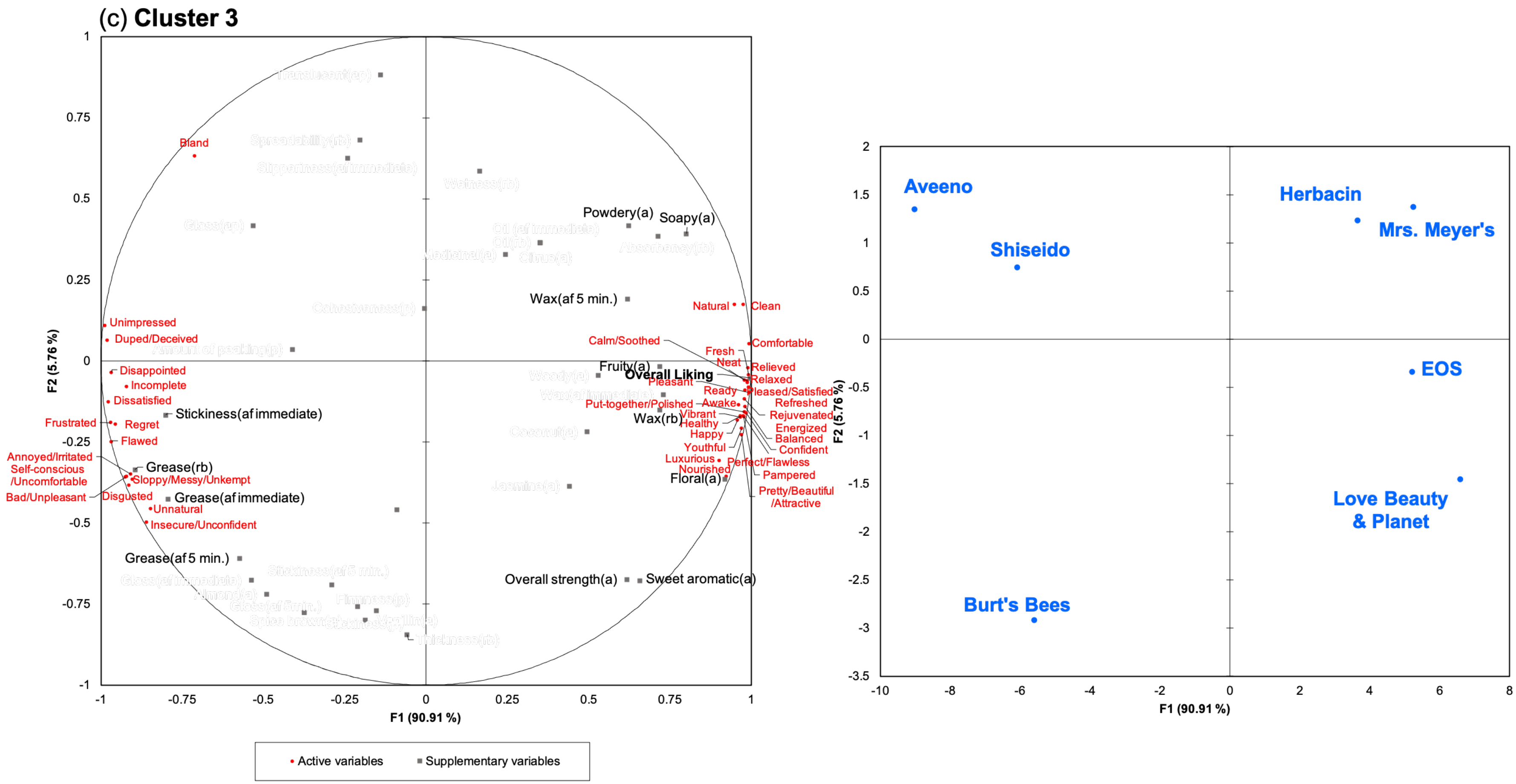


Figure 3.7. Emotion maps for the three clusters represented by the first two dimensions of PCA on mean emotional responses to the seven hand creams. Sensory descriptive data and consumer overall liking scores were projected into the maps as supplement variables. The maps only showed sensory attributes that had significant correlations (r >0.7, significantly different from 0, p<0.05) with any of the emotional terms.

The Relationship between Sensory Analysis and Consumer Perception for Different Consumer Segments

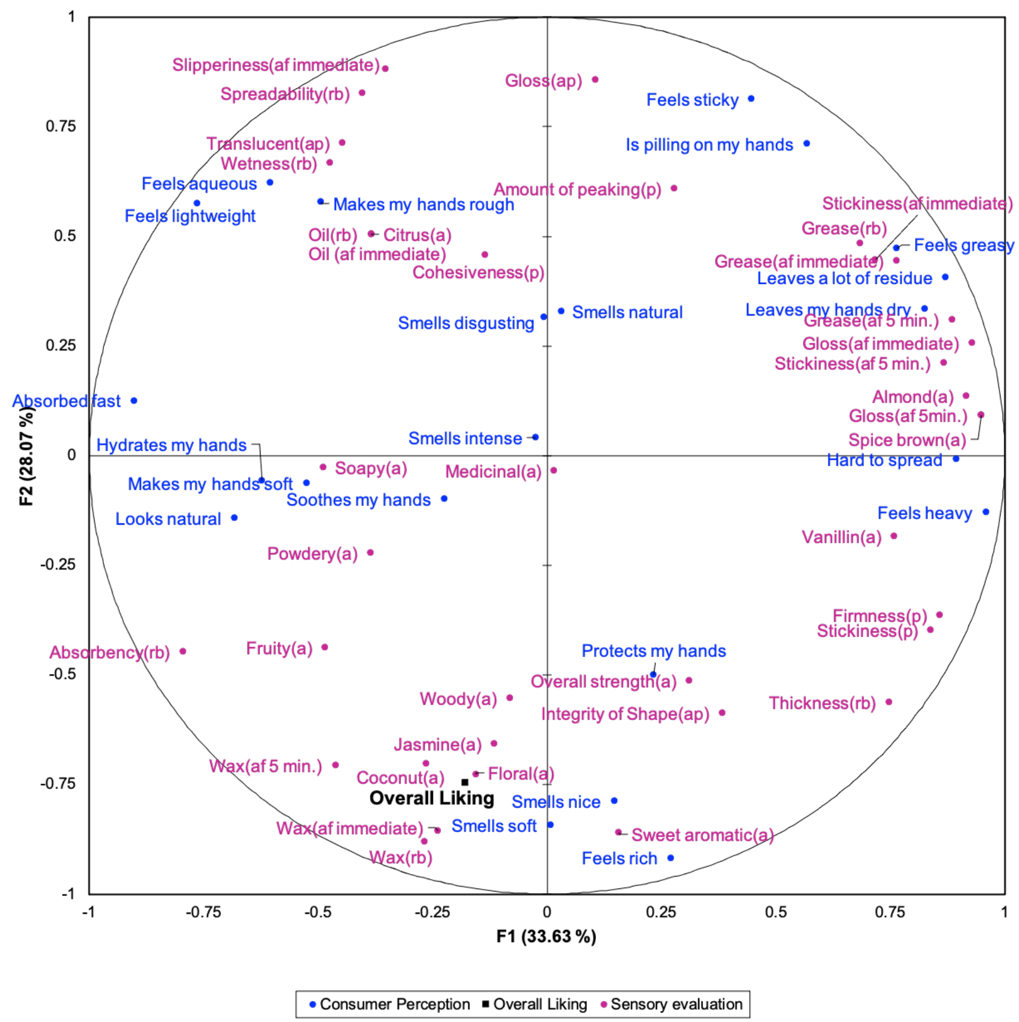
We were able to identify the sensory drivers of emotions for different consumer segments of liking patterns. To better explain the differences of sensory and emotion association between different consumer segments, the relationship between descriptive sensory data and consumer CATA data of the three clusters were explored.

MFA was conducted to compare the product profiles created by descriptive panel and consumer CATA for each consumer cluster. Figure 3.8 shows the variable correlation maps obtained from MFA. Consensus was found between consumers and the descriptive panel: all three clusters described greasy and sticky texture identified from the descriptive analysis in the same way as ‘feels greasy’ and ‘leaves a lot of residue’. The three consumer clusters also all perceived wet, spreadable, less thick texture, and translucent appearance as ‘aqueous’ and ‘lightweight’. However, the wet, spreadable, less thick texture and translucent appearance were only perceived as desirable for cluster 2, who related these attributes to ‘hydrating their hands’ and ‘looking natural’. Consumers of cluster 1, who were the thick and waxy texture likers, described the wet, spreadable, less thick texture to ‘make their hands rough’ and ‘not protect their hands’. This explained the reasons why these attributes elicited several negative emotions such as duped/deceived dissatisfied and flawed for cluster 1. In addition, the descriptive attributes waxy texture and afterfeel were highly correlated with consumer overall liking and the perception of ‘feeling rich’ in texture for cluster 1. This was probably why this texture elicited emotions such as refreshed/rejuvenated/energized and vibrant for this group. Different from cluster 1, cluster 3 considered the waxy texture and afterfeel to soothe and protect their hands, explaining the reasons why these attributes made them feel luxurious, pampered, and neat.

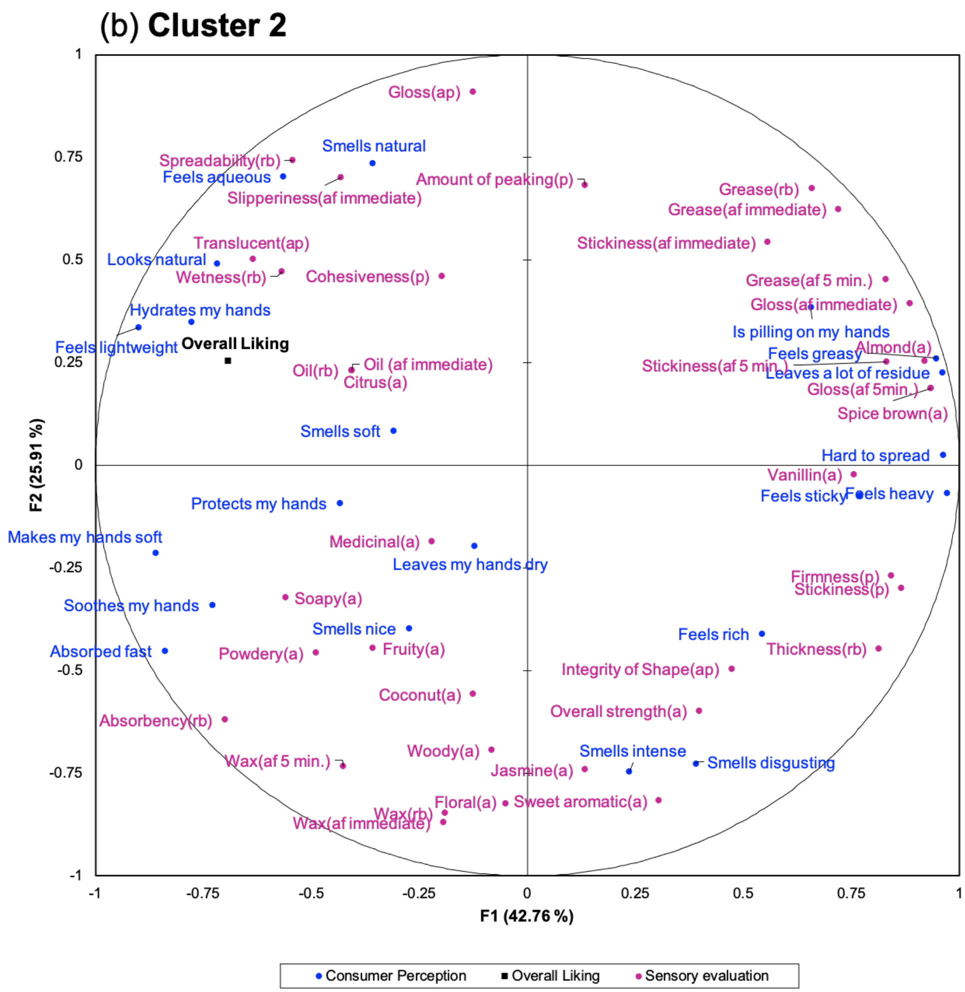
For aroma characterization, there were strong correlations between overall strength by descriptive analysis and ‘smells intense’ by consumer perception for both cluster 2 and cluster 3. ‘Smells intense’ was positively correlated with ‘smells disgusting’ and negatively correlated with overall liking for cluster 2. While for cluster 3, the intense-aroma-driven group, ‘smells intense’ was positively correlated with ‘smells nice’ and overall liking. Different from cluster 2 and cluster 3, consumers of cluster 1 only considered the aromas they disliked (soapy and medicinal) as ‘smells intense’.

In general, the different consumer segments did not differ largely on their perception of sensory attributes, as they could reach consensus on the perception of greasy, sticky, thin or thick, wet/aqueous texture and strong aroma. What differed was the values each cluster attached to the sensory characteristics of the hand creams depending on their preferences as well as the product efficacy for each individual depending on individual skin characteristics, while we did not find a significant difference in self-reported skin type between the three consumer segments. Cluster 2 described lightweight and aqueous texture would hydrate their hands, while cluster 1 felt this texture made their hands rough and could not protect their hands. This also explained the differences in sensory and emotion associations between consumer segmentations as consumers tended to associate the sensory attributes desirable for them with positive emotions, and undesirable attributes with negative emotions. Similar relationships between emotion, liking and consume perception have been observed in several food and beverage products (Pierguidi et al., 2019, 2020; Spinelli et al., 2019).

(a) **Cluster 1**



Slipperiness(af immediate)
Spreadability(rb)
Translucent(ap)
Wetness(rb)
Feels aqueous
Feels lightweight
Makes my hands rough
Oil(rb)
Citrus(a)
Oil (af immediate)
Cohesiveness(p)
Gloss(ap)
Feels sticky
Is pilling on my hands
Amount of peaking(p)
Stickiness(af immediate)
Grease(rb)
Feels greasy
Grease(af immediate)
Leaves a lot of residue
Smells disgusting
Smells natural
Leaves my hands dry
Grease(af 5 min.)
Gloss(af immediate)
Stickiness(af 5 min.)
Absorbed fast
Almond(a)
Gloss(af 5min.)
Spice brown(a)
Hard to spread
Hydrates my hands
Smells intense
Soapy(a)
Medicinal(a)
Makes my hands soft
Soothes my hands
Looks natural
Feels heavy
Vanillin(a)
Powdery(a)
Firmness(p)
Stickiness(p)
Absorbency(rb)
Fruity(a)
Protects my hands
Overall strength(a)
Woody(a)
Integrity of Shape(ap)
Thickness(rb)
Jasmine(a)
Wax(af 5 min.)
Coconut(a)
Floral(a)
Overall Liking
Smells nice
Wax(af immediate)
Smells soft
Sweet aromatic(a)
Wax(rb)
Feels rich
F2 (28.07 %)
F1 (33.63 %)
1
0.75
0.5
0.25
0
-0.25
-0.5
-0.75
-1
Consumer Perception
Overall Liking
Sensory evaluation



(b) Cluster 2
F2 (25.91 %)
F1 (42.76 %)
Gloss(ap)
Smells natural
Spreadability(rb)
Feels aqueous
Amount of peaking(p)
Slipperiness(af immediate)
Grease(rb)
Grease(af immediate)
Translucent(ap)
Stickiness(af immediate)
Looks natural
Wetness(rb)
Cohesiveness(p)
Grease(af 5 min.)
Gloss(af immediate)
Hydrates my hands
Is pilling on my hands
Feels lightweight
Overall Liking
Feels greasy
Almond(a)
Oil(rb)
Oil (af immediate)
Citrus(a)
Stickiness(af 5 min.)
Leaves a lot of residue
Gloss(af 5min.)
Spice brown(a)
Smells soft
Hard to spread
Vanillin(a)
Feels sticky
Feels heavy
Protects my hands
Makes my hands soft
Medicinal(a)
Leaves my hands dry
Firmness(p)
Stickiness(p)
Soothes my hands
Soapy(a)
Smells nice
Feels rich
Absorbed fast
Powdery(a)
Fruity(a)
Thickness(rb)
Integrity of Shape(ap)
Coconut(a)
Overall strength(a)
Absorbency(rb)
Woody(a)
Smells intense
Wax(af 5 min.)
Smells disgusting
Jasmine(a)
Floral(a)
Sweet aromatic(a)
Wax(rb)
Wax(af immediate)
Consumer Perception
Overall Liking
Sensory evaluation

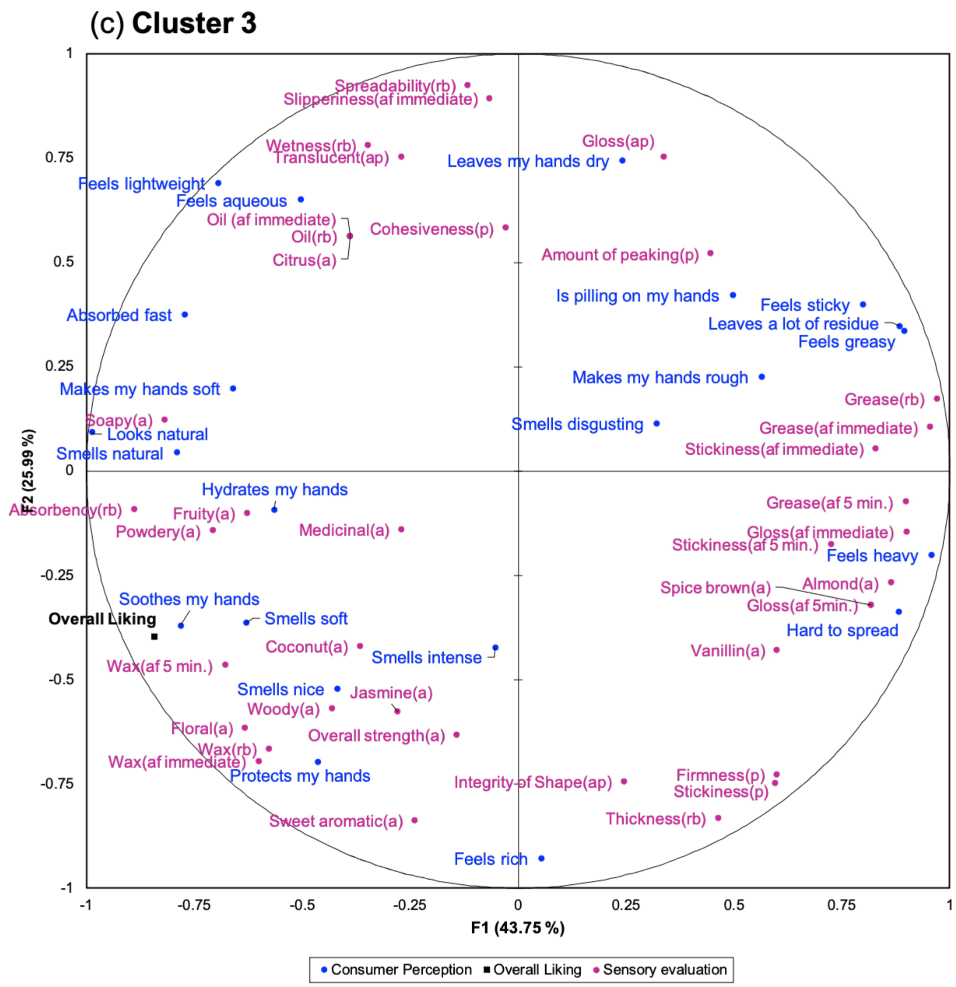


Figure 3.8. MFA variable correlation maps for each cluster obtained using descriptive analysis data and consumer CATA data with overall liking score as supplement variable.

Conclusion

In conclusion, this study identified the sensory drivers of liking as well as sensory-emotion associations for hand creams. Three consumer segments were identified according to their product liking patterns: cluster 1-the thick & waxy texture likers, cluster 2- mild scent & low-medium thickness likers, cluster 3-strong-scent likers. Sensory and emotion association was moderated by liking patterns and it differed across consumer segments. It was found across the three consumer clusters that 1) both aroma and texture attributes drove product-elicit emotions; 2) high intensities of certain aroma attributes seemed to elicit high-arousal emotions; 3) consumers related their desirable sensory attributes with positive emotions but undesirable attributes with negative emotions. Furthermore, comparing descriptive analysis data with consumer CATA data indicated the consumer segments differed more in the values they attached to sensory attributes than the sensory perception of attributes by themselves. The findings of this study could guide the development of new hand creams products targeting at different consumer segments. This study also demonstrated the importance of considering individual differences in liking when investigating sensory-emotion associations.

The limitation of this research is that the small number of hand creams are not enough to cover a wider combination of aroma profiles and texture profiles. In addition, a mix of diverse aroma profiles with texture profiles makes it hard to separate the effects of individual attribute in liking and emotions. Future works could utilize design of experiment (DOE) to control aroma and texture parameters to study how individual modality or attribute and their combinations affect emotional responses. Moreover, the emotion lexicon used in this study appear to be very unidimensional, the ratings of which are all highly correlated with overall liking scores. Even though product-elicited emotions provided further interpretation on consumers' overall liking patterns, they did not show superior product differentiating ability than the measure of overall liking. Tailoring this emotion lexicon specific to the hand cream category in the future with a smaller number of terms would contribute to a better emotion space for this product category.

References

ASTM Standard E1490-19. (2019). *Guide for Two Sensory Descriptive Analysis Approaches for Skin Creams and Lotions*. ASTM International.

Bhumiratana, N., Adhikari, K., & Chambers, E. (2014). The development of an emotion lexicon for the coffee drinking experience. *Food Research International*, *61*, 83–92.

Boinbaser, L., Parente, M. E., Castura, J. C., & Ares, G. (2015). Dynamic sensory characterization of cosmetic creams during application using Temporal Check-All-That-Apply (TCATA) questions. *Food Quality and Preference*, *45*, 33–40.

Chambers, E. (2018). Consensus Methods for Descriptive Analysis. In S. E. Kemp, J. Hort, & T. Hollowood (Eds.), *Descriptive Analysis in Sensory Evaluation* (pp. 211–236). John Wiley & Sons, Ltd.

Chrea, C., Grandjean, D., Delplanque, S., Cayeux, I., Le Calve, B., Aymard, L., Velazco, M. I., Sander, D., & Scherer, K. R. (2008). Mapping the Semantic Space for the Subjective Experience of Emotional Responses to Odors. *Chemical Senses*, *34*(1), 49–62.

David, O. A., David, D. O., Mogoase, C., Popescu, L. C., Giosan, C., & Pellegrino, A. (2019). Psychological effects and brain correlates of a rose-based scented cosmetic cream. *Journal of Sensory Studies*, *34*(6), e12536.

Di Donfrancesco, B., Koppel, K., & Chambers, E. (2012). An Initial Lexicon for Sensory Properties of Dry Dog Food: Sensory Lexicon for Dog Food. *Journal of Sensory Studies*, *27*(6), 498–510.

Guest, S., Dessirier, J. M., Mehrabyan, A., McGlone, F., Essick, G., Gescheider, G., Fontana, A., Xiong, R., Ackerley, R., & Blot, K. (2011). The development and validation of sensory and emotional scales of touch perception. *Attention, Perception, & Psychophysics*, *73*(2), 531–550.

Mora, M., Dupas de Matos, A., Fernández-Ruiz, V., Briz, T., & Chaya, C. (2020). Comparison of methods to develop an emotional lexicon of wine: Conventional vs rapid-method approach. *Food Quality and Preference*, *83*, 103920.

Painchault, T., Perrin, L., Loijens, L. W. S., Linden, X. van der, Theuws, H. M. M., & Zimmerman, P. H. (2020). Application of scientific measures to demonstrate the relaxing properties of Peony fragrance in hair care products. *Journal of Sensory Studies*, *35*(2), e12560.

Parente, M. E., Gámbaro, A., & Ares, G. (2008). Sensory Characterization of Emollients. *Journal of Sensory Studies*, *23*(2), 149–161.

Parente, M. E., Manzoni, A. V., & Ares, G. (2011). EXTERNAL PREFERENCE MAPPING OF COMMERCIAL ANTIAGING CREAMS BASED ON CONSUMERS' RESPONSES TO A CHECK-ALL-THAT-APPLY QUESTION: PREFERENCE MAPPING OF

ANTIAGING CREAMS BASED ON A CATA QUESTION. *Journal of Sensory Studies*, *26*(2), 158–166.

Pierguidi, L., Spinelli, S., Dinnella, C., Prescott, J., & Monteleone, E. (2019). Individual differences in perceived complexity are associated with different affective responses to alcoholic cocktails. *Food Quality and Preference*, *76*, 47–59.

Monteleone, E. (2020). Liking patterns moderate the relationship between sensory, emotional and context appropriateness profiles: Evidences from a Global Profile study on alcoholic cocktails. *Food Quality and Preference*, *83*, 103904.

Porcherot, C., Delplanque, S., Raviot-Derrien, S., Calvé, B. L., Chrea, C., Gaudreau, N., & Cayeux, I. (2010). How do you feel when you smell this? Optimization of a verbal measurement of odor-elicited emotions. *Food Quality and Preference*, *21*(8), 938–947.

Sanchez, K., & Chambers, E. (2015). How Does Product Preparation Affect Sensory Properties? An Example with Coffee: PREPARATION IMPACT ON COFFEE SENSORY PROPERTIES. *Journal of Sensory Studies*, *30*(6), 499–511.

Sanderson, T., & Hollowood, T. (2017). Temporal Methods for Assessment of Household and Personal Care Products. In J. Hort, S. E. Kemp, & T. Hollowood (Eds.), *Time-Dependent Measures of Perception in Sensory Evaluation* (pp. 362–387). John Wiley & Sons, Ltd.

Schouteten, J. J., De Pelsmaeker, S., Juvinal, J., Lagast, S., Dewettinck, K., & Gellynck, X. (2018). Influence of sensory attributes on consumers' emotions and hedonic liking of chocolate. *British Food Journal*, *120*(7), 1489–1503. https://doi.org/10.1108/BFJ-08-2017-0436

Sharma, C., Chambers, E., Jayanty, S. S., Sathuvalli Rajakalyan, V., Holm, D. G., & Talavera, M. (2020). Development of a lexicon to describe the sensory characteristics of a wide variety of potato cultivars. *Journal of Sensory Studies*, *35*(4).

Spinelli, S., Monteleone, E., Ares, G., & Jaeger, S. R. (2019). Sensory drivers of product-elicited emotions are moderated by liking: Insights from consumer segmentation. *Food Quality and Preference*, *78*, 103725.

Vázquez-Araújo, L., Nuncio-Jáuregui, P. N., Cherdchu, P., Hernández, F., Chambers, E., & Carbonell-Barrachina, Á. A. (2014). Physicochemical and descriptive sensory characterization of Spanish pomegranates: Aptitudes for processing and fresh consumption. *International Journal of Food Science & Technology*, *49*(7), 1663–1672.

Wortel, V. A. L., & Wiechers, J. W. (2000). Skin sensory performance of individual personal care ingredients and marketed personal care products. *Food Quality and Preference*, *11*(1–2), 121–127.

Xing, H., Vaught, C., & Chambers, E. (2020). Using multiple data analysis methods to guide makeup remover wipe optimization in a design of experiments consumer home use test. *Journal of Sensory Studies*, *35*(1).

Chapter 4 - Exploring Consumer Product Experience for Hand Creams Using Text Mining on Online Product Reviews

Abstract

This chapter explored consumer experience for hand cream products from the "voice of consumers"-online product reviews, aiming to investigate how information extracted from online reviews could facilitate sensory and consumer research. A total of 17,581 reviews representing 46 hand creams of different brands, price points, and sensory attributes were collected from Amazon and Ulta Beauty using a scraping software-Parsehub. Text analysis including text pre-processing, tokenization, term frequency calculation, topic modeling and sentimental analysis were performed primarily using the Tidytext package in R. Topic modeling using Latent Dirichlet allocation (LDA) identified five major topics consumers mentioned in these online reviews: greasiness & residue of the product, scent/fragrances of the product, skin feel & efficacy of the product, consumers' skin issues, and occasions when to apply the product. In addition, term frequency–inverse document frequency (tf-idf) was calculated for each rating group to identify the most relevant words and bigrams. It was found unpleasant scent and overall dissatisfied quality such as counterfeit product were the main reasons why consumers gave a rating lower than 4 stars. High efficacy such as moisturizing, and healing cracked skin, not greasy and smooth/soft skin feel were the drivers for a 5-star rating. These findings highlighted the importance of sensory experience and perception of efficacy in consumers' whole product experience. Consumer terminology regarding aroma, texture and skin feel of hand creams were also collected in this study.

Introduction

The latest few years have been characterized by a rapid shift to online shopping for the beauty care category. More and more consumers are purchasing their skin, hair, color cosmetic products through online retailers. Online platforms including retailers (e.g Amazon, Ulta Beauty and Sephora) and review websites (e.g Allure and MakeupAlley) have provided consumers a place where they can freely communicate about their product experience with other consumers as well as the companies. Online reviews are an important source of information that significantly influences consumer choice and the potential sales of a product (Dellarocas et al., 2007; Qi et al., 2016). Moreover, user-generated information can be a valuable source of insights guiding new product development and optimization. Conventional methods for data collection in consumer research can be time-consuming, labor-intensive, and costly. Meanwhile, most of these practices in sensory and consumer research place respondents in artificial settings that require them to think about their behaviors and adopt an analytical mindset (Köster, 2009). In comparison, online reviews are generated freely without the intervention of researchers and could provide richer information within shorter time and at lower cost. This makes it a great source of data for terminology development. For instance, Hamilton & Lahne (2020) developed a flavor wheel for whiskey from natural language processing of online review data of whiskey.

The rapid development data mining and natural language processing (NLP) techniques has enabled researchers to analyze massive amounts of unstructured text data collected from online sources such as webpages and bulletin boards. Text mining is an exploratory process that transforms unstructured text data to structured format so that patterns or insights can be explored (Han et al., 2010). The typical tasks include concept extraction, document summarization, text categorization and clustering, sentiment analysis, predicting models, etc. Many techniques and methods have been developed for text mining. Term frequency, topic modeling, and sentiment analysis are the frequently used methods to extract the key features of a category of products and possible product-elicited emotions. The application of online text mining techniques to analyze online review data has shown its potential in facilitating sensory and consumer research on personal care products. For instance, Kim & Kang (2018) collected product reviews for 42 Korean beauty balm BB creams and 194 competing BB creams from other countries from MakeupAlley, a cosmetic review site. By performing ratio analysis of words, Latent Semantic Analysis (LSA), Labeled Latent Dirichlet Allocation (L-LDA), and polarity analysis on the review data, the authors extracted 40 attributes that differentiate the Korean BB creams from BB creams of foreign countries (Kim & Kang, 2018).

This chapter was designed as an exploratory study, aiming to investigate the potentials of text mining on online reviews in understanding consumer experience for hand creams. The specific objectives were 1) to identify the main topics hand cream purchasers talk about in online reviews; 2) to extract key product features that drive high product ratings; 3) to try to explore consumers' emotions expressed in reviews of scented products vs non-scented products; 4) to build consumer terminology used to describe the sensory characteristics of hand creams.

Materials and Methods

Online Reviews

A total of 17,581 online product reviews were scraped from two websites Amazon (82.1%) and Ulta Beauty (17.9%) using a web scraping software-Parsehub (Toronto, Ontario, Canada). For reviews from Amazon, only those written by verified purchasers were collected. Each review includes information about product comment and product rating. All reviews use a 1–5-star rating system. These reviews represent a total of 46 hand creams of different brands, price points, sensory attributes. In addition, reviews in the collection cover scented products (63.0%) and scent-free products (37.0%).

Text Pre-processing

Data were all handled in the tidy text format (Wickham, 2014) in R (R Core Team, 2020). Two packages were mainly used in preparing and exploring the data: tidyverse (Wickham et al., 2019) and tidytext (Silge & Robinson, 2016). To prepare data for word frequency calculation and topic modeling, first, data cleaning was conducted on the review level. Reviews with low numbers of words (<=10) were dropped as they barely provided detailed insights about one's product experience. Second, all reviews were tokenized into unigrams/words and bigrams (pairs of consecutive words) separately. Bigrams might capture structure which cannot be identified when only counting single words and may provide context that makes tokens more understandable. Stop words such as "the", "of", "to" that had no added values to the main topics of the reviews were removed. The tidytext package considers "not" as a stop word. However, to make our tokens more meaningful, we did not remove "not" from the bigram tokens. In addition, we created a list of words for removal which were either too dominant words (e.g. hand(s), cream(s), lotion (s)) that prevented from finding interesting subdomains in our text, or words that did not contribute much meaning (e.g Amazon, review(s) and brand names). Furthermore, as the per-bigram counts appeared to be very sparse, stemming was done on words in the bigrams, so that bigrams of the same stems can be combined. Stemming means trying to make the tokens more generic by replacing it by a more generic form (Bouchet-Valat, M., & Bouchet-Valat, 2015). With stemming, the last part of the word is cut off to bring it to its more generic form. The unigram dataset was not treated by the stemming process as we wanted to keep the original feature of the words as much as possible.

Term Frequency

Ratings of the online reviews for hand creams were found to be highly skewed to the right (Figure 4.1). As a result, we consolidated all reviews that received ratings from 1-3 stars to be one single group for investigation. Three groups of rating levels were compared in total: the <4-star group, 4-star group, 5-star group. Frequencies of bigrams were calculated for each of the three rating groups, which were then plotted in world clouds for visualization. To examine how important a word was to a rating group in a collection of three groups, frequencies of unigram and bigram were adjusted for how rarely it was used in all groups by calculating term frequency-inverse document frequency (tf-idf). Tf-idf is normally calculated by multiplying the frequency of a term in a document (tf) with inverse document frequency (idf) (Silge & Robinson, 2017). Idf is denoted as the formula below (Silge & Robinson, 2017).

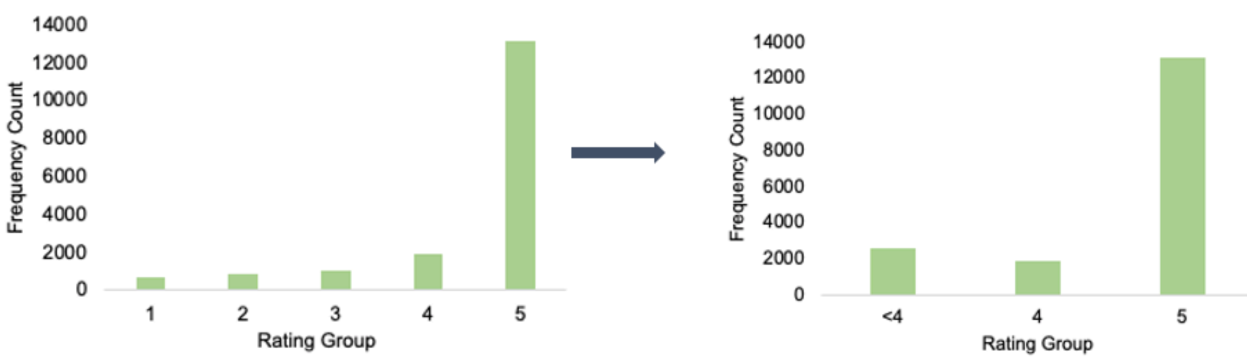


Figure 4.1. Distribution of reviews across rating groups. Reviews of 1-3 stars were consolidated into a single group for word frequency analysis

$$df\ (term) = \ln(\frac{n(documents)}{n(documents\ containing\ term)})$$

Topic Modeling

Topic modeling was performed on the whole dataset using Latent Dirichlet allocation (LDA) to detect the main themes or latent topics in the text data. Latent Dirichlet allocation

(LDA) is a particularly popular method for fitting a topic model (Blei et al., 2003). LDA treats each document as a mixture of topics, and each topic as a mixture of words, which allows documents to "overlap" each other in terms of content, rather than being separated into discrete groups, in a way that mirrors typical use of natural language (Blei et al., 2003; Silge & Robinson, 2017). LDA is a 'bag of words' technique, meaning the order of the tokens are not considered when performing the analysis. As a result, it lacks the consideration of tokens located in the text that are next to each other, which might miss crucial information of word collocations. To address this problem, both unigrams and bigrams were submitted to topic modeling. In addition, to make the document-term-matrix less sparse, only frequently used unigrams and bigrams were considered in the LDA model (occurred 5 times or more in the prepared data).

Sentiment Analysis

Dictionary-based 'bag-of-words' approach was used to identify the emotions/attitudes expressed in the online reviews for scented products (40 products, 11,064 reviews) and non-scented products (6 products, 6,517 reviews). Sentiment analysis in this study utilized the NRC Word-Emotion Association Lexicon (http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm) developed by Mohammad & Turney (2010, 2013). This lexicon contains a list of 14,182 English words and their associations with positive and negative sentiments, and eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust). Word-emotion associations included in this sentiment lexicon were initially constructed through manually annotating thousands of English words by crowdsourcing using Amazon Mechanical Turk (Mohammad & Turney, 2013). These word-emotion associations identified using the NRC lexicon don't necessarily mean denotations, instead, they are referred to as connotations, or implicit emotions (Mohammad, 2020). With sentiment analysis using the NRC lexicon, online review text was considered as a combination of individual words (corpus); the words in the corpus were matched against the sentiment lexicon so that they could be categorized in a binary fashion ('yes' or 'no') into positive, negative, anger, fear, anticipation, trust, surprise, sadness, joy and disgust. This analysis returned percentages of words in each document that had associations with each category of positive and negative sentiment and basic emotions. It should be noted that the NRC may categorize a word simultaneously to several categories. For example, the word 'love' is associated with both positive sentiment and joy; the word 'painful' is associated with negative sentiment, the emotions of sadness, fear, disgust, and anger; the word 'feeling' is categorized into both positive and negative sentiments. As a result, before conducting sentiment analysis, we manually examined the emotion associations from NRC lexicon with the most frequent terms in the hand cream online reviews and removed the entries that were not suitable in our case from the lexicon. In addition to sentiment analysis, the scented group was compared with the non-scented group for mean overall ratings by performing two sample t-test. A significant level of 0.05 was considered.

Results & Discussions

Main Topics in Hand Cream Reviews

After several iterations on the parameters of the topic model, a model with 5 topics was selected to be the most appropriate one to represent the whole dataset. Overall, topic modeling on unigrams and bigrams tokenized from online reviews of hand cream indicated 5 major topics consumers discussed in their online reviews. A token is considered important for a topic when it has a high predicted probability of occurrence within that topic. The most important (top 15) tokens per topic were shown in Figure 4.2. Each of the five topics were assigned with a theme it represented based on its most important tokens. 'Greasy', 'feel', 'absorbed', 'quickly' were the most frequently occurred tokens in topic 1, indicating this topic described the greasiness and residue of the product. Topic 2 mainly described the scent/aroma of the products. The most frequently occurred tokens in topic 3 were ‘day’, ‘winter’, ‘night’ indicating time, and ‘wash’, ‘cracked’ indicating occasions. Consequently, this topic represents the description of when to apply the product. Topic 4 consisting of terms such as ‘feel’, ‘smooth’, ‘moisturizing’, indicated the skinfeel and efficacy of the products. Topic 5 appeared to describe any skin issues of consumers as it had terms such as ‘dry’, ‘cracked’, ‘dry skin’ appearing frequently within this topic.

Sensory attributes such as greasiness, residue, skinfeel, and aroma have been reported to be important modalities affecting consumers’ likings in conventional sensory and consumer studies (Parente et al., 2011). This analysis has uncovered richer insights suggesting consumers’ skin issues and the related product efficacy, as well as the context for product usage seemed to be equally important factors that consumers considered when evaluating their whole product experience.

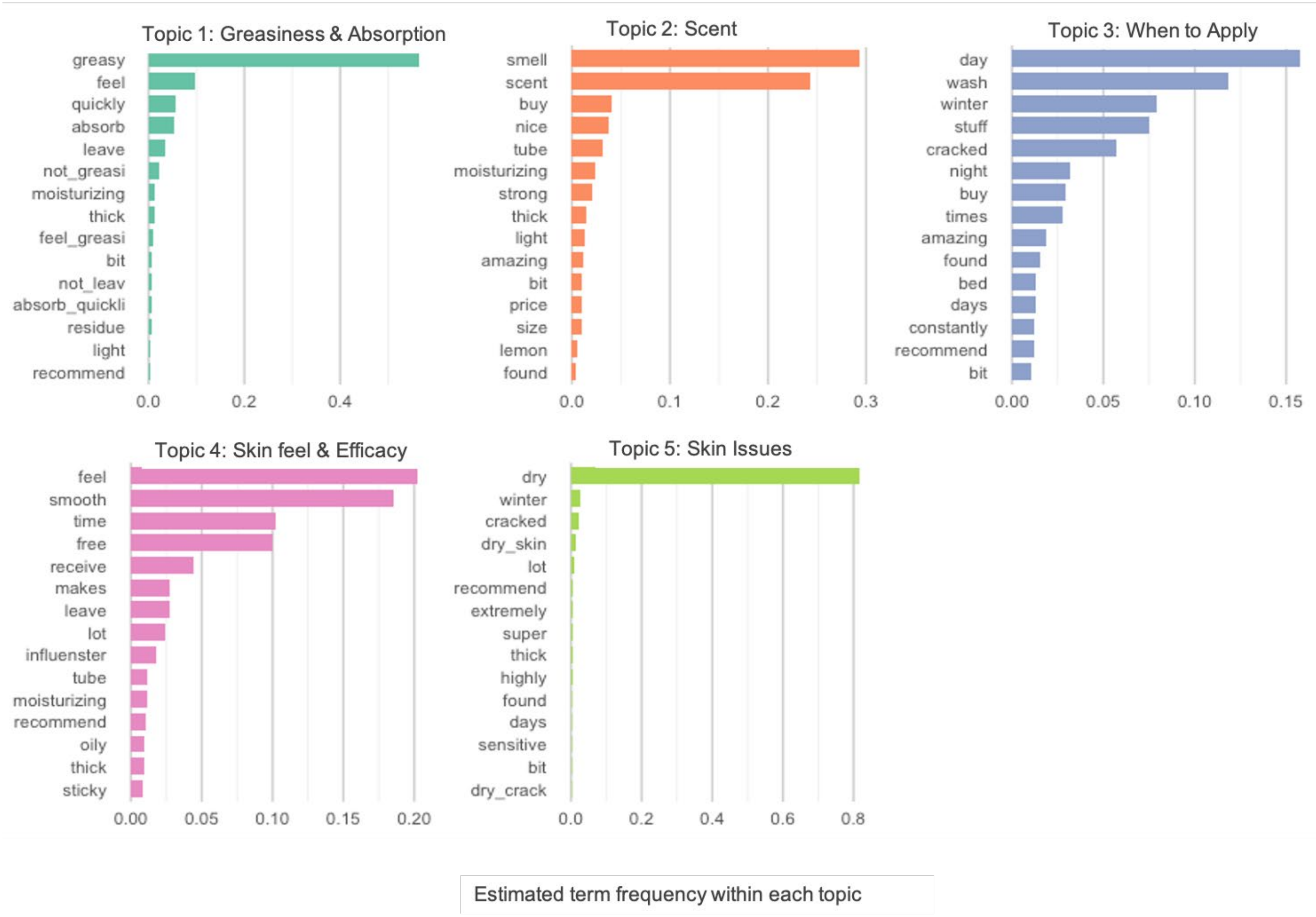


Figure 4.2. Five main topics identified from online reviews of hand creams by LDA topic modeling

Discriminating Features among Rating Groups

Frequencies of bigrams in each rating group were first visualized in Figure 4.3. The most frequently occurred bigrams in 5-star reviews were 'not greasy', 'dry skin', and 'highly recommend'. The most common bigrams in 4-star reviews were 'not greasy', 'dry skin' and 'feel greasy'. Differently for the <4-star reviews, the most frequently appeared bigrams in this group were 'not buy', 'dry skin' and 'not smell'. These results indicated that greasiness was an important factor affecting consumers' product experience for hand creams. The most desirable product attributes were found to be not being greasy, not leaving a lot of residues, being absorbed quickly. The undesirable product attributes were identified as not having pleasant smell, being greasy, not being absorbed and not moisturizing.

Figure 4.3. Word clouds based on the frequencies of unigrams and bigrams for each rating group

Calculating frequencies of bigrams could uncover some trends across different rating groups, however, it was hard to extract discriminating features between rating groups, especially in cases when there were terms occurring frequently in all rating groups (e.g. ‘dry skin’ and ‘not greasy’). To address this issue, term frequency-inverse document frequency (tf-idf) were calculated using all unigrams and bigrams for each rating group within the context of three rating groups. Tf-idf identifies the specific terms more unique to each rating group, hence, would help detect the discriminating terms among rating groups. Figure 4.4 showed the tokens (unigrams and bigrams) of the highest tf-idf for each rating group. These figures suggested the possible reasons why a consumer would rate a product a score of less than 4 stars, 4 stars or 5 stars. It was found that the main reasons for a rating less than 4 star could be due to the dissatisfaction of the overall quality of the product (related terms such as trash, broken, counterfeit, unsealed, watered, half empty), and unpleasant scent considered by a consumer (related terms such as horrible smell, awful smell, not pleasant). A 5-star rating was given because consumers perceived the high efficacy of the products (related terms such as lifesaver, solved, improved), desirable skinfeel (related terms such as soft smooth, not greasy) and scent. Terms of high tf_idf values identified in the 4-star group were a mix of positive and negative descriptors of the products. A 4-star indicated that consumers were satisfied with the products overall, however, they had a hope of improvement in some aspects of the product experience such as the texture (related terms such as slightly greasy, pretty light) or the price (related terms such a bit pricy). In general, the results of tf-idf for terms in each rating groups suggested that perceived efficacy and desirable skinfeel were the key for a consumer to be satisfied with a product. Unpleasant scent and perceived low quality were the main drivers for a low rating below 4 stars.

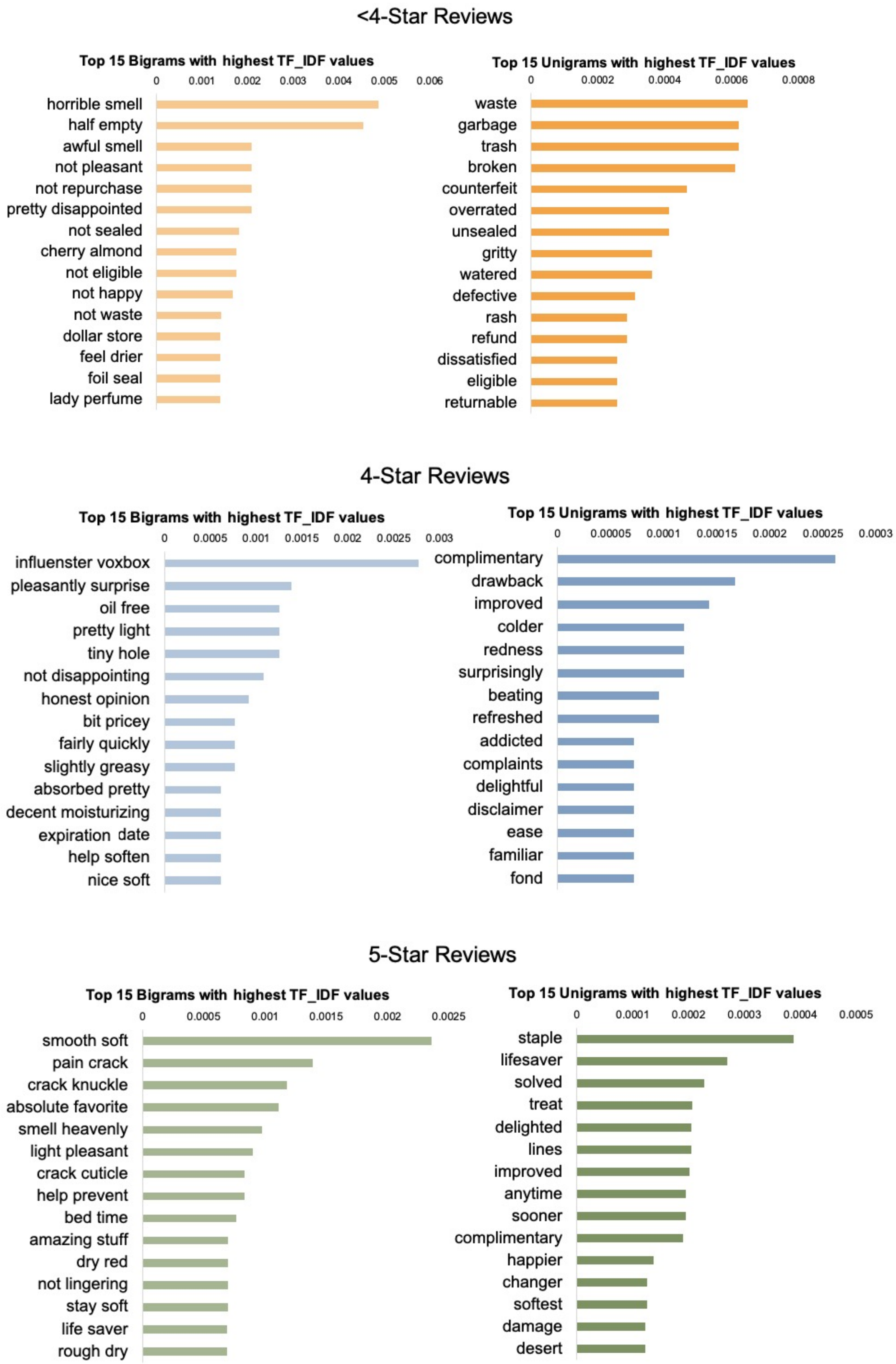


Figure 4.4. Top 15 bigrams and unigrams with highest tf-idf values in each rating group

Consumer Terminology for Sensory Attributes

Topic modeling and tf-idf highlighted the importance of sensory perception in consumers' whole product experience. As a result, the language consumers used to describe the scent, texture and skinfeel of hand creams was extracted by identifying the preceding words and following words of 'feel', 'feeling', 'smell', 'scent', 'fragrance' in all bigrams. The terminology extracted from online reviews (Table 4.1) can be used in future consumer studies for product characterization purposes.

Table 4.1. Consumer terminology used to describe the sensory experience of hand creams

Aroma			**Texture**	**Skinfeel**
acidic	faint	overpowering	balmy	chapped
almond	floral	peach	chalky	cracked
aloe	flowery	perfumey	creamy	dry
banana	fruity	pleasant	gel	hydrated
cake	grapefruit	pomegranate	greasy	itchy
candy	harsh	powdery	oily	moisturized
chamomile	heavy	rose	rich	silky
chemical	herbal	sharp	slimy	slippery
cherry	honey	spicy	thick	smooth
chocolatey	lavender	strong	velvety	soft
citrusy	lemon	sweet	watery	sticky
coconut	masculine	tart	waxy	supple
cookie	medicinal	tea		tacky
cucumber	menthol	unpleasant		wet

eucalyptus	non-perfumey	vanilla		

Sentiment Analysis

Two sample t-test performed on overall ratings (Figure 4.5) suggested that the scented group of products was rated at parity with the non-scented group ($p>0.05$). Interestingly, sentiment analysis using the NRC lexicon performed on the reviews of scented products and non-scented products captured the differences in word-emotion associations for the two groups of products. Compared to reviews of the scent-free product group, higher percentages of positive and joy words (e.g. love, wonderful, perfect, pleasant, happy) were used in the reviews of scented products (Figure 4.6). These results suggested that scented products may be more likely to elicit positive emotions than scent-free products. It was also found that more negative, fear, sadness and anger words were used in the reviews of the non-scented products than those of scented products (Figure 4.6). However, it did not necessarily mean that non-scented products were more likely to elicit negative emotions as the most frequent negative words in the reviews of scent-free products were 'cracked', 'cracking' and 'crack'. These words were used by consumers to express their current skin issues and explain the reasons why they bought the hand creams, rather than to describe their perception and feelings after using the hand creams.

One should note that sentiment analysis of text data is different from conventional measurement of emotions. The idea of sentiment analysis is to uncover the emotional tone expressed in the text data. Sentiment analysis assumes that words convey affect not only explicitly through their core meaning but also implicitly through connotation (Mohammad, 2020). Conventional emotion lexicons focused on a small number of terms that denote emotions, for instance, sad denotes sadness; or happy denotes joy. Lexicons used in sentiment analysis, on the other hand, tend to focus on a larger set of words that are associated with or connotate an emotion. Table 4.2 showed the most common words (top 10) in the reviews of hand cream products that were identified to connote positive and negative sentiments and the eight basic emotions based on the NRC lexicon. Some of these associations might seem arbitrary in the case of hand cream reviews: for example, ‘job’ was categorized as a word with positive sentiment; ‘sweet’ was categorized to convey surprise. This was because the word-emotion associations defined by the NRC lexicon were fixed and not domain dependent. Domain-specific sentiment lexicon should be developed and utilized if future studies aim to uncover emotions from the text data of a specific category of product. In addition, the sentiment analysis in this study was based on the ‘bag-of-word’ model, which may lose information about grammar and the context of usage of each word.

Table 4.2. Most common words in the reviews of hand cream that contributed to positive and negative sentiments and different basic emotion categories according to the NRC lexicon.

Positive	**Negative**	**Joy**	**Trust**	**Surprise**
love	cracked	love	recommend	wonderful
recommend	cracking	perfect	perfect	gift
perfect	bad	wonderful	wonderful	pleasant
wonderful	crack	favorite	favorite	sweet
favorite	disappointed	gift	pretty	deal
worth	overpowering	pretty	constantly	miracle
job	smelling	excellent	excellent	surprised
gift	bleeding	happy	happy	excited
pretty	painful	pleasant	pleasant	expect
excellent	hate	clean	effective	hope
Anticipation	**Sadness**	**Fear**	**Disgust**	**Anger**
perfect	bad	cracked	greasy	cracked
gift	disappointed	bad	sticky	bad
pretty	bleeding	bleeding	bad	disappointed
happy	painful	painful	disappointed	painful
pleasant	hate	hate	smelling	hate
daily	awful	awful	bleeding	awful
expected	terrible	terrible	painful	terrible

healing	worse	worse	hate	crazy
glad	crazy	crazy	weird	hurt
continue	hurt	hurt	awful	horrible

The NRC lexicon may associate a word with several emotion categories.

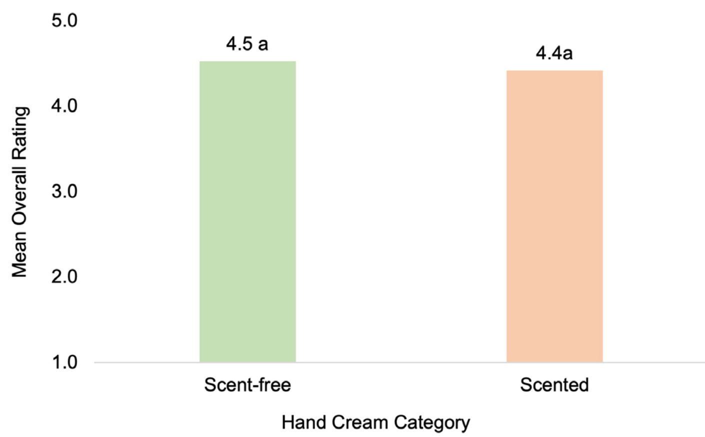


Figure 4.5. Mean overall ratings of scent-free products and scented products. Mean overall ratings with the same letter designation were not statistically different on a 1–5-star scale (alpha=0.05)

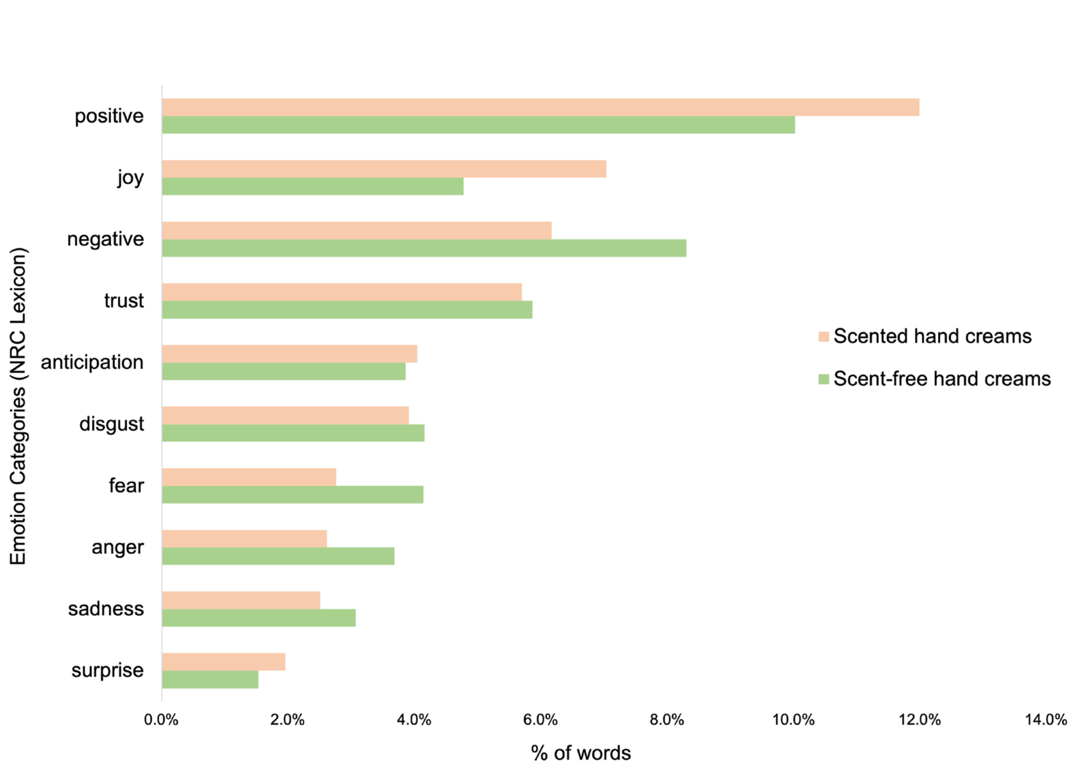


Figure 4.6. Emotion words expressed in the reviews of scented products vs non-scented products

Conclusion

This study demonstrated an exploration process in understanding consumer reviews of hand creams. Consumer terminology regarding aroma, texture and skinfeel of hand creams were also collected. The results of word frequency, tf-idf analysis, and topic modeling highlighted the importance of sensory experience, perception of efficacy, context of usage in consumers' whole product experience. Perceived efficacy and desirable skinfeel such as making hand smooth were the key reasons for a 5-star rating. Poor overall quality and unpleasant smell led to a rating below 4 stars. Comparing the findings from online review mining in this study with those from external preference mapping (EPM) in Chapter 3, both methods identified greasy and sticky skinfeel to be the major sensory attributes that drove consumer disliking, and high absorbency and low residue to be the attributes driving consumer liking. The advantages of conventional research using EPM were that sensory drivers of liking could be examined in detail for different consumer segments. Online review mining in this study was unable to identify consumer segments, however, it provided rich information about whole product experience from consumers which could be used to improve the design of conventional sensory and consumer research. For instance, online review mining of hand creams suggested the importance of efficacy in consumer experience of the hand cream category. In future studies, instead of generating preference mapping for skincare products, efficacy mapping such as a hydrating mapping seems to be more useful as it can help identify the sensory attributes that are related the efficacy of the products.

There were several limitations. Firstly, this study was conducted based on online reviews; demographic information of the reviewers was not collected, preventing us from getting insights from a targeted population compared to traditional sampling methods. Secondly, most of the online reviews were positive reviews with overall ratings larger than 4 stars. As a result, the analyses performed in this study were on a rating level without considering the differences between products. The 46 samples had a wide range of number of reviews and all products were rated well with no large differences on overall consumer likings, making it hard to extract the key features that differentiate on a product level. Further, topic modeling and sentiment analysis performed in this study were all based on the bag-of-word model, in which a sentence was represented by a bag of words without considering the order of the words in a sentence. Information loss regarding the order of words and the grammar in a sentence might decrease the precise of the insights uncovered by topic modeling and sentiment analysis. Future analysis may address this issue using more sophisticated models such as Word2vec to understand and predict the meaning of a word considering the surrounding words. The last limitation of this research was that the authenticity of the review data was not examined. Insights generated from fake reviews might provide misleading information. There have been some commercial customer review analytical tools which identify fake reviews by examining the reviewers' profiles and review contents. Future research could utilize this type of tools to filter out fake reviews before performing detailed text analysis.

References

Bouchet-Valat, M., & Bouchet-Valat, M. M. (2015). Package 'SnowballC'.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. the Journal of machine Learning research, 3, 993-1022.

Dellarocas, C., Zhang, X. (Michael), & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, *21*(4), 23–45.

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimedia Tools and Applications, 78(11), 15169-15211.

Hamilton, L. M., & Lahne, J. (2020). Fast and automated sensory analysis: Using natural language processing for descriptive lexicon development. *Food Quality and Preference*, *83*, 103926.

Han, H., Back, K.-J., & Barrett, B. (2010). A consumption emotion measurement development: A full-service restaurant setting. *The Service Industries Journal*, *30*(2), 299–320.

Kim, S. G., & Kang, J. (2018). Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews. *Information Processing & Management*, *54*(6), 938–957.

Köster, E. P. (2009). Diversity in the determinants of food choice: A psychological perspective. *Food Quality and Preference*, *20*(2), 70–82. https://doi.org/10.1016/j.foodqual.2007.11.002

Mohammad, S. M., & Turney, P. D. (2010). *Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon*. 9.

Mohammad, S. M., & Turney, P. D. (2013). CROWDSOURCING A WORD-EMOTION ASSOCIATION LEXICON. *Computational Intelligence*, *29*(3), 436–465.

Mohammad, S. M. (2020). Practical and ethical considerations in the effective use of emotion and sentiment lexicons. arXiv preprint arXiv:2011.03492.

Parente, M. E., Manzoni, A. V., & Ares, G. (2011). EXTERNAL PREFERENCE MAPPING OF COMMERCIAL ANTIAGING CREAMS BASED ON CONSUMERS' RESPONSES TO A CHECK-ALL-THAT-APPLY QUESTION: PREFERENCE MAPPING OF ANTIAGING CREAMS BASED ON A CATA QUESTION. *Journal of Sensory Studies*, *26*(2), 158–166.

Qi, J., Zhang, Z., Jeon, S., & Zhou, Y. (2016). Mining customer requirements from online reviews: A product improvement perspective. *Information & Management*, *53*(8), 951–963.

Silge, J., & Robinson, D. (2016). tidytext: Text mining and analysis using tidy data principles in R. Journal of Open Source Software, 1(3), 37.

Silge, J., & Robinson, D. (2017). Text mining with R: A tidy approach. " O'Reilly Media, Inc.".

Wickham, H. (2014). Tidy data. Journal of statistical software, 59(1), 1-23.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., ... & Yutani, H. (2019). Welcome to the Tidyverse. Journal of open source software, 4(43), 1686.

Appendix A - Candidate Emotion Terms for Lexicon Development (324 Terms in Total)

acceptance	awake*	concerned	dirty	flattered	guilty	irritable	old*	rave	sexy	threatened	vicious
accomplished	awe	confident	disappointed	flawed	happy	irritated	optimistic	ready	shame	thrilled	vitality
active	awesome	confused	discomfort	flawless	harsh	joyful	outrageous	reassurance	shiny	tired	vulnerable
addicted	awful	content	discouraged	fond	hassle	keen	outstanding	rebellious	shocked	tolerant	warm
adoration	awkward	cool	disgusted	fooled	hate	lazy	overwhelmed	refreshed	shy	torture	weird
adult*	bad	courageous	dislike	foolish	healing	loath	painful	regret	sickened	tough	well-being
adventurous	balanced	courtesy	dismay	frantic	healthy	lonely	pale	rejoicing	silly	traumatized	wild
afraid	bashful	cracked	dissatisfied	free	heartbroken	love	pampered	rejuvenated	sincere	tricked	wonder
aggravated	bland	craving	dull	fresh	hesitation	lovely	panic	relaxed	skeptical	trust	wonderful
aggressive	blessed	crazy	duped	friendly	honest	loyal	paranoid	relief	sloppy	ugly	worried
agony	bliss	creative	eager	frizzy*	hooked	lucky	passionate	replenished	soft	unattractive	wow
alert	bold	creepy	ecstatic	frumpy*	hope	luxurious	patient	respectful	soothed	uncomfortable	youthful
allure	boosted	cruelty	effortless	frustrated	hopeless	manic	peaceful	responsible	sophisticated	unconfident*	
amazed	bored	crumbling	elegant	fun	horrible	mature	perfect	restored	special	understanding	
amusement	bothered	crushed	embarrassed	funky	hot	merit	pleasant	revitalized	steady	unfortunate	
angry	bright	crusty	energized	furious	hurt	messy	pleasantly surprised	rewarded	stimulated	unhappy	
annoyed	bulky	curious	enhanced	fussy	hysteria	miserable	pleased	ridiculous	strange	unhealthy	
anxious	calm	cute	enjoyment	fuzzy	immaculate	motivated	polished	risky	stressed	unimpressed	
apologetic	camouflage	damaged	enthusiastic	generous	impatient	naked*	polite	romantic	strong	unique	
appealing	carefree	daring	envious	gentle	impressed	nasty	popular	rotten	stubborn	unkepmt*	
appreciation	careless	decadent	excellent	genuine	incomplete	natural	powerful	rough	stunned	unnatural	
apprehensive	caring	deceived	excited	glad	incredible	nausea	prestigious	rude	stupid	unpleasant	
approachable*	cautious	decent	exposed	glamorous	indifferent	neat	pretty/beautiful	sad	stylish	unpleasantly surprised	
ashamed	charming	defeated	failure	glitter	indulgent	nervous	professional	satisfied	suffering	unprofessional*	
assured	cheerful	delighted	faithful	glowing	inferior	nightmare	promising	scared	superior	unruly	
astonished	chic	depressed	fanatic	good	insecure	nourishment	protected	screwed	supple	unsure	
attentive	clean	desire	fantastic	gorgeous	inspired	obsessed	proud	secure	suspicious	upset	

attracted	clueless	desperation	fascinated	grateful	intimidated	odd	put-together*	self-conscious*	tender	vain	
attractive	comfortable	determined	flaky	grimy	invigorated	offended	quiet	sensitive	terrible	vanity	
avid	complaint	diligent	flashy	gross	irresponsible	offensive	radiant	sensual	terrific	vibrant	

*These 12 terms were not sourced from online reviews but were gathered from Talavera and Sasse (2019).

Appendix B - Aroma Attributes, Definitions, and References in Descriptive Analysis

Attribute	Definition	Reference
Overall strength	The overall impact of the product aroma.	Shiseido Hand Cream=2.0, Nivea Crème= 10.0
Soapy	A pungent, slightly fragrant aromatic with fatty base notes characteristic of unscented hand soap.	Ivory Soap Bar = 7.5
Floral	Sweet, light, slightly fragrant aromatic associated with fresh flowers.	Welch's White Grape Juice = 4.0
Jasmine	An intense, slightly pungent, sweet, floral aromatic with underlying green, musty/dusty notes.	Silver Cloud Jasmine Extract = 8.5
Fruity	Sweet, light, slightly fragrant aromatic associated with fruit.	Welch's White Grape Juice = 6.0
Citrus	The citric, sour, astringent, slightly sweet, peely, and somewhat floral aromatics which may include lemons, limes, grapefruits, and oranges.	Peels of Lemon and Lime = 4.5
Almond	Light, sweet, nutty aromatic reminiscent of almonds.	McCormick Almond Extract = 7.5
Coconut	The slightly sweet, nutty, somewhat woody aromatic associated with coconut.	McCormick Coconut extract = 7.5
Spice brown	The aromatics commonly associated with brown spices, may include spice such as cinnamon, cloves, nutmeg, all spice, and others.	McCormick Ground Allspice=5.5
Sweet aromatics	An aromatic associated with the impression of a sweet substance.	Lorna Doone Cookie=4.5
Vanillin	An extremely sweet non-natural aromatic often associated with vanilla, cotton candy and marshmallows.	Le Nuz de Vin #40 = 6.0
Medicinal	A clean, sterile aromatic characteristic of antiseptic like products such as Band-Aids, alcohol, and iodine.	Listerine Original (Diluted) = 8.0
Powdery	The fragrance effect produced by the interaction of long-lasting mossy, woody, sweet, and crystalline elements. Many perfumes leave a powdery overall impression after evaporation of the fresh and floral ingredients.	Johnson and Johnson Baby Powder = 8.0
Woody	The sweet, brown, musty, dark aromatics associated with a bark of a tree.	Wood chips = 4.0
Leather	Aromatic impression that is characterized as being somewhat damp, dark, and heavy, and is connected to musty, new leather.	Le Nuz de Vin #45 = 4.0

Appendix C - Demographics of Consumer Segments in the Home Use Test

Age	Overall	Cluster 1 (N= 39)	Cluster 2 (N=27)	Cluster 3 (N=34)
18-24 years	9%	8%	7%	12%
25-34 years	31%	31%	22%	38%
35-44 years	29%	33%	33%	21%
45-54 years	31%	28%	37%	29%
Ethnicity				
Caucasian/White	77%	82%	81%	68%
African American/Black	7%	8%	4%	9%
Hispanic/Latino	5%	3%	7%	6%
Asian-All countries	9%	8%	7%	12%
Other	2%	0%	0%	6%
Household Income				
Under $25,000	2%	0%	0%	6%
$25,000 to $34,999	3%	5%	0%	3%
$35,000 to $49,999	10%	10%	7%	12%
$50,000 to $59,999	11%	10%	11%	12%
$60,000 to $69,999	8%	5%	11%	9%
$70,000 to $99,999	21%	23%	15%	24%
$100,000 to $149,999	26%	23%	33%	24%
$150,000 or more	19%	23%	22%	12%
Education				
Some high school	1%	3%	0%	0%
High school degree	2%	3%	0%	3%
Some college	15%	15%	15%	15%
College degree	39%	31%	52%	38%
Some graduate studies	8%	5%	7%	12%
Graduate degree	35%	44%	26%	32%
Skin Type				
Dry	22%	21%	26%	21%
Normal-dry	62%	67%	59%	59%
Normal	13%	8%	15%	18%
Combination	3%	5%	0%	3%
Hand Cream Usage Frequency				
Once a day	9%	5%	11%	12%
2-3 times a day	60%	69%	52%	56%
4-5 times a day	17%	15%	19%	18%
6-7 times a day	11%	8%	15%	12%

More than 7 times a day	3%	3%	4%	3%

Appendix D - Two-Way ANOVA for Overall Liking, Attribute Liking and Attribute Intensity in the Home Use Test

	Sample Effect		Cluster Effect		Sample*Cluster Effect	
	F	p-value	F	p-value	F	p-value
Overall Liking	14.54	**< 0.0001**	0.94	0.391	16.59	**< 0.0001**
Appearance Liking	7.91	**< 0.0001**	2.54	0.079	3.60	**< 0.0001**
Aroma Liking	8.16	**< 0.0001**	4.73	**0.009**	8.68	**< 0.0001**
Texture Liking	15.43	**< 0.0001**	0.09	0.912	5.07	**< 0.0001**
Afterfeel Liking	17.93	**< 0.0001**	1.00	0.367	5.85	**< 0.0001**
Aroma Intensity	71.74	**< 0.0001**	19.05	**< 0.0001**	2.60	**0.002**
Thickness Intensity	66.32	**< 0.0001**	8.86	**0.000**	0.97	0.473
Greasiness Intensity	30.08	**< 0.0001**	0.32	0.725	2.12	**0.014**

Appendix E - Two-Way ANOVA for Emotion Responses in the Home Use Test

	Sample Effect		Cluster Effect		Sample*Cluster Effect	
	F	p-value	F	p-value	F	p-value
Awake (+)	8.54	**< 0.0001**	2.81	0.061	3.86	**< 0.0001**
Balanced (+)	4.49	**0.000**	3.46	**0.032**	4.94	**< 0.0001**
Calm/Soothed (+)	5.11	**< 0.0001**	2.69	0.069	8.27	**< 0.0001**
Clean (+)	5.65	**< 0.0001**	4.17	**0.016**	6.03	**< 0.0001**
Comfortable (+)	6.96	**< 0.0001**	4.26	**0.015**	7.90	**< 0.0001**
Confident (+)	8.08	**< 0.0001**	4.09	**0.017**	6.12	**< 0.0001**
Fresh (+)	9.80	**< 0.0001**	3.02	**0.049**	6.56	**< 0.0001**
Happy (+)	11.62	**< 0.0001**	4.19	**0.016**	8.78	**< 0.0001**
Healthy (+)	5.62	**< 0.0001**	1.60	0.202	5.31	**< 0.0001**
Luxurious (+)	9.58	**< 0.0001**	4.28	**0.014**	5.21	**< 0.0001**
Natural (+)	4.41	**0.000**	3.66	**0.026**	6.49	**< 0.0001**
Neat (+)	4.64	**0.000**	5.36	**0.005**	6.49	**< 0.0001**
Nourished (+)	3.87	**0.001**	2.02	0.133	5.92	**< 0.0001**
Pampered (+)	11.30	**< 0.0001**	4.29	**0.014**	6.68	**< 0.0001**
Perfect/Flawless (+)	6.16	**< 0.0001**	5.94	**0.003**	6.05	**< 0.0001**
Pleasant (+)	9.95	**< 0.0001**	5.90	**0.003**	9.83	**< 0.0001**
Pleased/Satisfied (+)	10.37	**< 0.0001**	2.95	**0.053**	9.88	**< 0.0001**
Pretty/Beautiful/Attractive (+)	7.12	**< 0.0001**	6.66	**0.001**	5.87	**< 0.0001**
Put-together/Polished (+)	5.14	**< 0.0001**	8.43	**0.000**	6.20	**< 0.0001**
Ready (+)	4.66	**0.000**	4.25	**0.015**	6.95	**< 0.0001**
Refreshed/Rejuvenated/Energized (+)	8.16	**< 0.0001**	3.05	**0.048**	6.56	**< 0.0001**
Relaxed (+)	6.29	**< 0.0001**	3.74	**0.024**	6.89	**< 0.0001**
Relieved (+)	3.92	**0.001**	2.88	0.057	5.22	**< 0.0001**
Vibrant (+)	8.31	**< 0.0001**	6.23	**0.002**	4.60	**< 0.0001**
Youthful (+)	5.21	**< 0.0001**	2.97	0.052	3.58	**< 0.0001**
Annoyed/Irritated (-)	5.62	**< 0.0001**	1.62	0.199	5.80	**< 0.0001**
Bad/Unpleasant (-)	6.40	**< 0.0001**	1.29	0.276	8.46	**< 0.0001**
Bland (-)	16.88	**< 0.0001**	7.49	**0.001**	4.28	**< 0.0001**
Disappointed (-)	7.88	**< 0.0001**	1.24	0.291	7.98	**< 0.0001**
Disgusted (-)	4.94	**< 0.0001**	0.75	0.472	5.10	**< 0.0001**
Duped/Deceived (-)	4.29	**0.000**	2.91	0.055	4.58	**< 0.0001**
Flawed (-)	5.03	**< 0.0001**	1.09	0.337	4.22	**< 0.0001**
Incomplete (-)	3.47	**0.002**	2.61	0.074	2.79	**0.001**
Insecure/Unconfident (-)	3.50	**0.002**	2.20	0.111	3.38	**< 0.0001**

Self-conscious/Uncomfortable (-)	4.55	**0.000**	3.33	**0.036**	4.61	**< 0.0001**
Sloppy/Messy/Unkempt (-)	7.81	**< 0.0001**	3.21	**0.041**	3.21	**0.000**
Unhealthy (-)	1.62	0.137	2.13	0.120	1.88	**0.033**
Unimpressed (-)	9.08	**< 0.0001**	2.45	0.087	8.00	**< 0.0001**
Unnatural (-)	3.51	**0.002**	1.32	0.269	2.19	**0.011**
Dissatisfied (-)	6.54	**< 0.0001**	1.99	0.138	7.01	**< 0.0001**
Frustrated (-)	3.64	**0.001**	2.20	0.112	5.43	**< 0.0001**
Regret (-)	1.67	0.127	1.35	0.259	4.37	**< 0.0001**