5-2013

Measuring Arousal through Physiological Responses to Packaging Designs: Investigating the Validity of Electrodermal Activity as a Measure of Arousal in a Realistic Shopping Environment

Daniel Hutcherson
Clemson University, danhutcherson@gmail.com

MEASURING AROUSAL THROUGH PHYSIOLOGICAL RESPONSES TO PACKAGING DESIGNS: INVESTIGATING THE VALIDITY OF ELECTRODERMAL ACTIVITY AS A MEASURE OF AROUSAL IN A REALISTIC SHOPPING ENVIRONMENT

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Packaging Science

---

by
Daniel Edward Hutcherson
May 2013

Accepted by:
Dr. R. Andrew Hurley, Committee Chair
Dr. Chip Tonkin
Dr. Shaundra B. Daily
Dr. Andrew T. Duchowski

ABSTRACT

The validity of electrodermal activity is investigated in the context of packaging design evaluation as a sensitive measure of arousal. Analysts agree that purchasing decisions are subconscious and emotional decisions, contrasted against the popular belief that consumers make purely rational decisions. To understand the personal and rapid character of a consumer's purchasing decisions we must find methods of which to measure and interpret consumer reactions to various packaging designs. Focus groups are discussed as antiquated research methods and new, advanced technologies are outlined as physiological responses. Past literature displays methodological approaches to using electrodermal activity measures in consumer studies, however these studies observe participants outside of the shopping context.

This thesis argues that the behavior found in a traditional lab setting is not indicative of true consumer behavior experienced in the shopping context. Through utilizing electrodermal activity within the realistic shopping environment, designers can attempt to better collect and interpret a consumer's preference towards a packaging design. This physiological approach to packaging design evaluation proposes a methodology for collecting and interpreting electrodermal activity to better match the consumer behavior discussed by marketing professionals.

The findings of an additional post-pilot study indicate that, at this time, there is not enough research available to disregard electrodermal activity as a tool for packaging designers. Significance is found with the use of the Self Assessment Manikin as an aid to qualify the use of the emotional stimuli presented. No significance is found in the

reporting of electrodermal activity and eye tracking data from the full-scale study. While a shopping context will produce natural, subconscious reactions to packaging, the quick and non-linear movements of consumers make recording and understanding these reactions very difficult. Before a complete dissociation between packaging design evaluation and electrodermal activity is formed, it is recommended that researchers closely parallel the fields of neuroscience, psychology, and physiology in order to understand the benefits of updated measures of emotion. A visual concept is introduced as a method of triangulating measures from the exports of electrodermal activity, eye tracking, and self-report measures. Parameters are recorded throughout this thesis to document a foundation for future research in the field of consumer emotions, highlighting the struggles and successes of measuring one's physiological responses. A general theme is established in hopes that it will endure with future research, stating that true reactions are captured in true contextual frames.

DEDICATION

To my Mom, Dad and entire family: thank you for your support and patience. Let's go to the beach!

ACKNOWLEDGMENTS

I would first like to thank my advisor, Dr. Andrew Hurley and my committee members, Dr. Chip Tonkin, Dr. Shaundra Daily, and Dr. Andrew Duchowski for all of their guidance and support. Thank you for your time and review of this research topic.

My appreciation is directed towards Mr. Harris A. Smith who has funded my research within this program. Thank you for providing me with this opportunity!

A *big* thank you is extended to the past and current students that form the graduate team, including: Rachel Randall, Katie Thackston, Josh Galvarino, Joanna Fischer, Toni Gomes, Andy Pham, Andrew Ouzts, Erin Synder, Nathan Bailey, Gabrielle Conlon, Wilson Sansbury, Felipe Fernandes, and Julianne Holmes. I have enjoyed working, studying, and traveling with you all! I wish for you all the best of luck and much future success. A special thanks is extended to Ouzts for his help and guidance throughout this program. I appreciate your advice and collaboration.

I would also like to thank the friendly Q-Sensor team at Affectiva for your voluntary support in quickly processing the data in this thesis. I couldn't have completed this thesis without your help.

With much excitement, I would like to extend my thanks to Will Keown for visually capturing our research in the most beautiful way. In addition to your time, I sincerely appreciate the level of craft and detail applied to these images.

Finally, I would like to thank all of those involved at the Sonoco Institute of Packaging Design and Graphics. This project would not have been possible without you.

TABLE OF CONTENTS

Table of Contents (Continued)

Page

Table of Contents (Continued)

Page

LIST OF TABLES

LIST OF FIGURES

List of Figures (Continued)

Figure Page

List of Figures (Continued)

List of Figures (Continued)

Figure Page

List of Figures (Continued)

Figure | Page

CHAPTER ONE

INTRODUCTION

"Emotion Leads to Action"

Consumers are faced with a challenge in today's retail environments; shopping tasks outline the strong presence of emotional design strategies. Purchase decisions are directed by emotional actions, expressing the unconscious motives of the human brain (Lindstrom, 2010). Brief phasic events, approximately 2.5 seconds in length, frame intervals of decision making for most consumers (Blakeslee, 2004; Witchalls, 2004). A history of marketing reflects the directive that "emotion leads to action." Indeed, studies have shown that 95% of thinking is unconsciously realized where consumers make purchasing decisions based on emotion rather than rational thought (Blakeslee, 2004; Lindstrom, 2010; Pop, Radomir, Ioana, & Maria, 2009). As a result, the design and development of most products in the marketplace attempt to utilize a person's emotional triggers (Pawle & Cooper, 2006; Thaler & Sustein, 2009). We are guided by packaging to make critical decisions affecting our health, happiness, and well being. Rob Girling, a principle designer at Artefact, reflects the statement that people are the products of thinking below the level of awareness. Surprisingly, decision-making is not as cumbersome a task, where reason is highly dependent on emotional value judgments. Fundamentally, emotion is shown to play a big role in decision making, however, as designers, there is no shared view on how to record, systemize, and reproduce the process of a successful design. This process of creation and introduction into the marketplace leaves packaging designers with an opaque process. The known outcome, an emotional response to the consumer, is clearly impactful, but hard to control (Girling, 2012). In the competitive marketplace of product duplicates, it is necessary for a package to capitalize on the emotional response of the consumer (Pawle & Cooper, 2006).

In attempting to record these emotions, researchers have largely relied on focus group sessions as a means of self-report measures to gain valuable insight into consumer perception in packaging designs (Folch-Lyon & Trost, 1981; Zaccai, 2012). However a shift between the consumer's conscious and subconscious directives occur through these traditional studies, strictly removing the consumer from the environment. Traditional methods of research in packaging must be updated to accurately measure a consumer's emotion in reaction to packaging designs in the physical shopping context.

Physiological Measures of Consumer Behavior

Emotional engagement can be used as an extremely reliable predictor of marketplace success (Pradeep, 2010). To discover more about the emotional dialogue between consumers and packaging, neurophysiological measures are captured to illustrate the effects of packaging on consumers' purchasing decisions (Chamberlain & Broderick, 2007). Self-report measures are used as a compliment to realistic data gathered from behavior recording devices. With the introduction of several tools, researchers are able to ask how a consumer feels in several reliable forms (Picard & Daily, 2005). Wireless and unobtrusive eye tracking glasses are used as tools to immerse consumers in realistic shopping environments, where real-life tasks enable consumers to react naturally to packaging designs (Hurley, 2011). These exchanges between technology and people help to record and annotate the quiet conversations between consumers and packaging.

The introduction of wireless recording devices in market research allows researchers to gather more realistic responses from consumers in the shopping context. Through utilizing direct physiologic means, such as electrodermal activity, eye tracking, and electroencephalography, a more descriptive and accurate evaluation of consumer preference is illustrated, forecasting higher purchasing decisions. With the support of these physiologic tools, future costly production runs in packaging can be saved by a descriptive means of consumer testing (Tonkin & Duchowski, 2011; Tonkin, Ouzts, & Duchowski, 2011).

This thesis is a discussion of two physiological devices: eye tracking (eye movements and areas of interest) with the addition of electrodermal activity (cognitive arousal). Electrodermal activity is introduced to the toolset of the package designer in order to investigate the validity of the tool when used inside the realistic shopping context.

Eye Tracking

Eye tracking is a useful technique in recording what consumers look at when shopping. Focusing on the consumer's field of view allows researchers to see where attention is diverted, forming descriptive regions of interest. Eye trackers can record the geographic location of fixations on a retail shelf in consumer studies. Head mounted, mobile eye tracking glasses, (Figure 5) provide a lightweight and unobtrusive means of monitoring eye movements (Chandon, 2002; Duchowski, 2007). Packaging studies reveal how longer fixation times correlate to increased sales or higher purchasing decisions (Hurley, Galvarino, Thackston, Ouzts, & Pham, 2012). Eye tracking is useful in documenting the geographic locations of eye movements, however, eye tracking does not reveal the emotional valence of the fixations (Duchowski, 2007).

Electrodermal Activity (EDA)

Balanced activity within the sympathetic and parasympathetic divisions of the autonomic nervous system supports the regulation of physiological states of arousal. The parasympathetic nervous system (PNS) and sympathetic nervous system (SNS) are often compared to the brakes and gas of a car respectively; helping "rest and digest" versus "fight and flight" activities. While the PNS dilates blood vessels leading to the digestive tract, stimulates salivary glands, and constricts the bronchioles of the lungs, the SNS prepares the body to act on changing environmental conditions by accelerating heart rate, constricting blood vessels, and raising blood pressure. When SNS activity increases, sympathetic fibers that surround eccrine sweat glands modulate the production of sweat. The skin, in turn, momentarily becomes a better conductor of electricity (i.e., electrodermal activity). This electrodermal activity can be measured as conductance (skin conductance) or resistance by different sensors (Dawson, Schell, & Filion, 2007; Picard & Daily, 2005).

The Validity of Electrodermal Activity in Packaging

To acknowledge the subconscious, emotional directives of the consumer, this thesis combines the efforts of eye tracking and electrodermal activity as physiological measures of consumer behavior. The validity of electrodermal activity, as a measure of arousal, is investigated with a thorough methodological approach applied inside the physical shopping context. The following review of literature discusses the shopping context as the missing link in traditional consumer studies using EDA and other physiologic tools. A pilot study and full-scale study are provided as goal oriented objectives to fully immerse and involve the consumer in the experience of shopping. The findings are evaluated to discuss the validity of the tool as (1) a measure of arousal towards packaging designs (2) within a realistic shopping context of both time and appearance.

CHAPTER TWO

REVIEW OF LITERATURE

Packaging Design

Packaging is the silent "salesman on the shelf" that sells products in a competitive environment (Rettie & Brewer, 2000). Successful package designs look to strike emotion, necessary as the quick decision maker (Silayoi & Speece, 2004). Packaging design is an important tool in the marketing of consumer goods because of its use in communicating the benefits of a product to the consumer (Rundh, 2009). Modern consumers are always looking for ways in which to spend less time shopping, specifically with food items. This is a key point of focus for packaging designers where one must design for the fast-paced shopper. The pressure of time inside the shopping environment reduces the time for detailed considerations (Warde, 1999). By today's standards, the package can be considered as an equal to the product (George, 2005). Consumers feel that the package is the product until the product is opened. Once open, the package is thrown away, recycled, or happily reused (Rundh, 2009).

The overall creativity and design of the package should strike an emotional appeal with the consumer, necessary to persuade someone into making a purchase decision (Rundh, 2009). Packaging designers aim to employ shapes and graphics in packaging designs to be emotionally appealing and compatible with a certain brand (Raghubir & Greenleaf, 2006; Wansink, 1996). Anthropomorphization, attributing human form to things non-human, is a technique that some designers and marketing professionals rely on to provide a close connection between the humanized consumer and the desirable product

(Landwehr, McGill, & Herrmann, 2011). Design factors in packaging are researched to gain consumer praise. These types of factors involved in the design process include the following: shape or form, size, orientation (up or down), alignment (left or right), color, color combinations, and imagery (Westerman et al., 2013). Specific to packaging design, vivid and attractive imagery is the most positive influence towards a successful packaging design (Silayoi & Speece, 2004).

In addition to emotionally attracting the consumer, the packaging design process follows these parameters towards a successful design element: package production (low cost, material sustainability, temperature control, atmospheric conditioning), product filling, distribution (minimum space and weight), retail outlet marketing (emotional and attractive design, product security), consumer purchase (convenient, easy, and safe), consumer storage (quality over time), and consumer usage (ease of use) (Rundh, 2009). The urgency of this thesis' narrative is shown with claims that consumer design preference changes over time due to trends and cultural effects (Carbon, 2010; Westerman et al., 2013). It is imperative to establish a more effective way of measuring reactions to designed objects in our world.

A Path to Action

"Emotion leads to action" (Harris, 2006). Emotional marketing is a term that Consoli states as the process by which large enterprises leverage their marketing operations to arouse the consumer with their brand (Consoli, 2009). Emotions, similar to the functionality of a product, influence the purchasing decision-making process to a high degree (Consoli, 2009) as they hold a strong impact on attention (Liao, Corsi, Lockshin, & Chrysochou, 2012). Scientists support the view that people find personal reward from internal feelings, which can be caused from intrinsic and extrinsic causes (Ahn & Picard, 2006). People tend to prefer things that cause a high level of pleasure. In a speech to the Promotion Marketing Association of America, Light spoke of the emotional attachment that consumers have with brands, or *trustmarks* as he describes them. The value of a certain brand can be best gauged by the value of the trust that the people have for the item (Light, 1993). Affective responses towards different stimuli is a major determinant of a product's overall likeness (Landwehr et al., 2011). People buy products because of the emotional payoff instead of the certain features or benefits of a certain product (Harris, 2006).

Kroeber-Riel (1979) has suggested that arousal specifically enhances the processing of advertising information (Kroeber-Riel, 1979). Singh and Churchill conclude further that arousal can be highly effective in increasing long-term memory (Singh & Churchill, 1987). Subconscious emotional responses to stimuli have been shown to nudge consumers into making purchase decisions with minor, sometimes zero, cognitive thinking (Shiv & Fedorikhin, 1999).

In addition to attracting attention, emotions help consumers in making purchase decisions. Attention, with added emotional feelings can lead to the acceptance of a message when provided to the consumer (Poels & Dewitte, 2006). For a brand to be considered by anyone in today's fast paced society, it must grab attention with strong, emotional messages. The successful use of emotions in directed messages can lead to automatically driving the action of the consumer (Liao et al., 2012). It is known in advertising (a form of packaging) that those ads that fail to cause an arousal will, in turn, have no effect on the consumer. Knowing this highlights the importance of testing the measure of arousal prior to expensive product launches in advertising, packaging, and all fields related to design. Products which have a desire to succeed in the marketplace will be constructed off of measures known to excite consumers repeatedly (Kroeber-riel, 1979).

Interdisciplinary and Interjectory Research

Dickson & Sawyer report that minor cognitive effort usually occurs when consumers stand in front of products on grocery shelves. Usually only a few seconds are spent deciding on a certain packaged good (Dickson & Sawyer, 1990). The lack of activity and the speed at which the consumer moves has caused many researchers to investigate new methods, interjecting information and processes from a wide range of fields. Recent research involving consumers and their emotional affective responses to packaging has relied on the tools of qualitative and self-report based quantitative methods of evaluation. Other fields, such as print and TV advertising, have leveraged the talents of other fields of research to gain more reliable insight into measuring emotions. Physiological research in packaging design is a fairly new field of inquiry. Packaging is learning from the far-reaching research attitudes of other marketing areas (Liao et al., 2012). Psychobiology is an example of one such interdisciplinary field of research combining biology, physiology, neurology, and psychology. Other fields have merged to create focuses termed neuromarketing, neuropsychology, and psychophysiology (Groeppel-Klein, 2005; Kroeber-riel, 1979).

An intersection of human behavior and the physiological processes of the nervous system (particularly the brain's activity) can be shared with the research needs of the packaging industry (Kroeber-riel, 1979). In understanding consumer reactions to advertising campaigns and supporting a further research and exploration of EDA research methods, LaBarbera argues that marketers clearly need better pretesting tools when guiding the design of product offers and communication devices. It is important to emotionally motivate consumers in order to get them to respond (LaBarbera & Tucciarone, 1995). Ferber calls researchers to action stating that we must reach to an "interdisciplinary research in consumer behavior." Introducing additional fields into the study of packaging designs on consumer emotions will enable further possibilities and revealing research opportunities (Ferber, 1977). Light claims that the time is right for new and different ways of thinking and acting in the field of marketing, a close connection to packaging design (Light, 1993).

Attitudes about Emotion Research

It is the purpose of this thesis to shed a positive light on the use of physiological measures to better understand human behavior in the shopping context. Through showing evidence from the acceptance of these measures in consumer behavior, we should also acknowledge that there are negative reactions towards this form of research.

It should be understood that physiological measures are an innocuous path towards better understanding human behavior for the benefit of society through health and happiness.

Emotional marketing is beneficial to businesses because of its success in creating a close and personal relationship with the customer in the marketplace. This close connection to the consumer provides the foundation for a competitive position on the market (Consoli, 2009). Marketing professionals advise that designers understand consumer goods as people, not products. A successful brand will engage the consumer much like a personal relationship (Harris, 2006).

Research involving neuro-psycho-physiological reactions to consumer goods has worried people into questioning if there is a "buy button" inside the brain for which designers and marketing professionals are searching (Blakeslee, 2004; Witchalls, 2004). A letter by Gary Ruskin was written to Senator John McCain, chairman of the Senate Commerce Committee, in 2004 to address fears of neuromarketing due to its implications on politics and public health. Ruskin expressed concern that marketing professionals could possibly discover a power to modify human behavior (Ruskin, 2004). To respond to these concerns, scientists have addressed these issues as "gross misunderstandings and distortions" of the power of brain imaging tools in consumer research (Blakeslee, 2004). Reiman, a researcher at Atlanta neuromarketing firm BrightHouse, expresses that the spirit behind this mode of research is aiming to find the consumer's heart, not the brain (Wahlberg, 2004).

Gilbert R. Lindberg, research director at Cunningham & Walsh and other agencies has supported the role that EDA has in predicting consumer behavior in the marketplace, by stating that "when we can predict, we can control" (Lindberg, 1988). While Lindberg is correct that EDA will provide sufficient evidences to predict consumer behavior, it is important to not control the consumer but to conversely "nudge" (Thaler & Sustein, 2009) the consumer towards making healthy and beneficial decisions. The word "nudge" is introduced by Thaler as an appropriate and healthy alternative to the negative connotations derivable from the word "control." The ethical standards of scientific research should support agendas regarded as humanistic rather than personal self-advancement in the eyes of large corporations. The future trends in EDA and other physiologic research in packaging should uphold narratives to aid and abide by the progressive health of the participants of our consumer driven society (LaBarbera & Tucciarone, 1995).

Methods of Evaluation in Packaging Design

Traditional (Self-Report)

Classical or traditional methods of measuring emotion have relied on self-report measures or questionnaires (Picard & Daily, 2005). Past studies involving emotion and packaging are defined with qualitative self-reporting measures (Gofman & Moskowitz, 2009; Nancarrow, Wright, & Brace, 1998; Rundh, 2009; Silayoi & Speece, 2004). Self-report measures, such as focus groups and surveys, are not a reliable source of consumer feedback because it is difficult for consumers to express their emotional reactions (Valentino-DeVires, 2010; Zaccai, 2012). Focus groups facilitate conscious decision-making, opposite to the real-life, unconscious decisions made in the shopping environment. Consequently, the consumer's unconscious experience is instantly falsified and misinterpreted (Sukhvinder, 2011).

Three examples of self-report measures can include think-aloud and thought-list procedures, descriptive Likert scales, and the Self-Assessment Manikin.

Interviews and think-aloud processes attempt to reveal the true feelings or emotions of a person in relation to a specific experience. This self-reporting technique requires the participant to report all personal thoughts during the processing of a task. This is a detailed process used to gather detailed information on mental processes. Techniques of verbal reports should be used carefully as they can easily produce invalid results (Shapiro, 1994). Retrospective Think Aloud (RTA) is a verbal method that is applicable as a supplementary self-reporting measuring to this thesis. RTA is a method that involves the verbalization of a participant's performance as they view a recording of their past eye movements. A usability study completed by Guan and colleagues assesses this technique as a valid measure of user performance (Guan, Lee, Cuddihy, & Ramey, 2006).

Likert scales are used to describe personal emotions and moods through descriptive words (Petrides & Furnham, 2000). Examples of words used in 5-7 point Likert scales include the following mood adjective: *pleasant, unpleasant, neutral, down, guilty, insecure, lonely, anxious, happy, cheerful, relaxed, and satisfied* (Myin-Germeys, Van Os, Schwartz, Stone, & Delespaul, 2001). These scales can prove to be invalid where personal and cultural differences across participants could create different connotations towards word choices present on the evaluation scales (Heine, Lehman, Peng, & Greenholtz, 2002; Wierzbicka, 1999).

In order to remove all textual information, the Self-Assessment Manikin scale was designed to measure emotional reaction to stimuli as a valid self-reporting measure. The S.A.M. scale is a non-verbal assessment that consists of pictorial information to directly measure the pleasure, arousal, and dominance of a participant in response to a particular stimulus (M. Bradley & Lang, 1994; Coan & Allen, 2007; Liao et al., 2012; Morris, 1995). An examples of the manikin figures used in the S.A.M. method can be viewed in Figure 1 (M. Bradley & Lang, 1994).

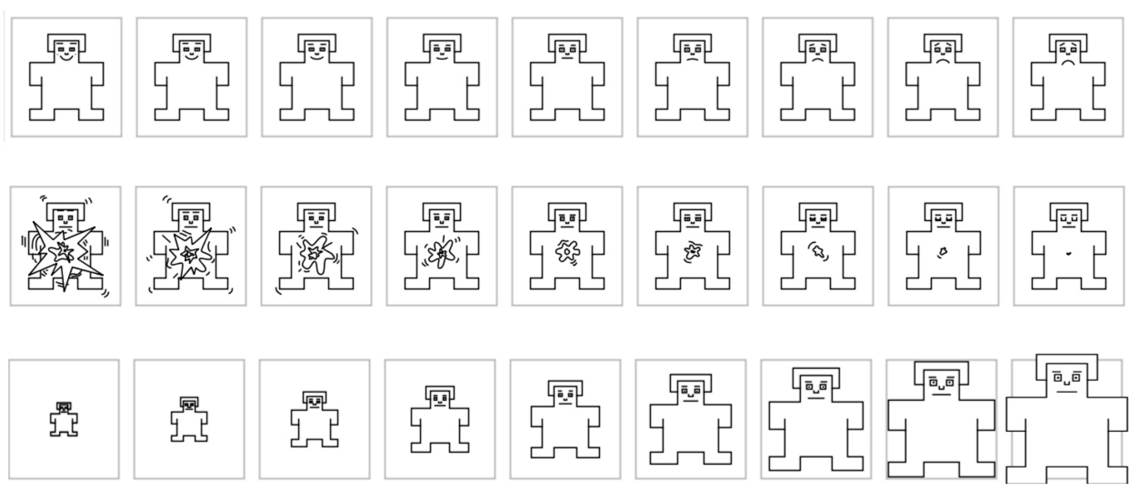


Figure 1: Self-Assessment Manikin Scale (S.A.M.). From Bradley and Lang, 1994, © Elsevier Science Ltd., UK, used with permission.

Self-report measures can be argued as non-valid measures of emotional responses due to the fact that verbalized responses are simply reflections of a past experience. It is fairly easy for a participant to misinterpret their emotions when reflecting on a past experience (Liao et al., 2012; Nisbett & Wilson, 1977). As an example, textual information is used to detect emotions towards a particular product being viewed and purchased: sentences are divided into tokens or words and analyzed with an algorithm. Affective vectors are added to reveal the emotional quality of the words used (Consoli, 2009). These traditional measures are retrospective in nature, involving a post behavior, self-interpretation (LaBarbera & Tucciarone, 1995). They require people to verbally interpret a past experience as a reflection. Sometimes people can find it difficult to accurately, precisely, and uniformly articulate emotional responses (Harris, 2006). Advertising research is limited when research methodologies are dependent on the verbal responses of consumers (LaBarbera & Tucciarone, 1995). In using attitudinal behavior to predict consumer behavior to different marketing stimuli, a poor, or false predictor of behavior is established (Light, 1993).

LaBarbera provides several of the misleading examples found when relying on self-report measures: participants will report what they think to be pleasing to the study interviewer, people will respond to aesthetic or entertainment value instead of motivational impact, and also consumers might intentionally attempt to remain consistent with the conclusions derived from a desired public self-image (LaBarbera & Tucciarone, 1995). Penn writes that previous research utilizing self-report measures only could be invalid where the responses can only be understood as "disguised or post-rationalized" reflections (Penn, 2006).

The reason why we cannot entirely rely on the self-report measures of consumers, is not because study participants are devious or untruthful, consumers are sincerely largely unaware of the triggers that really motivate them. Rational research questions will directly result in rational responses. Rationality does not exist in the impulse-shopping environment of the consumer. Consumers are led to make choices for many irrational and emotional reasons (LaBarbera & Tucciarone, 1995). A form of cognitive evaluation is required from the participant regarding a past experience. In some cases, due to subconscious processing, the person may not even be consciously aware of the arousal (Groeppel-Klein, 2005). A time lag is experienced from stimuli encounter to cognitive reaction (Groeppel-Klein, 2005). Normally, verbal measures produce different responses from participants. With verbal, self-reporting, a detour is directed across the direct measures of physiological processes. Self-reporting is only a person's conscious perception of an arousal, not always a complete or valid measure (Kroeber-riel, 1979). With the majority of thinking being unconsciously realized, how is it that researchers can rely on the conscious efforts of focus groups and other self-reporting measures (Pop et al., 2009)?

To find solutions to understanding consumer behavior, marketing teams have initiated war stories to understand purchase intent. Some researchers are skipping straight to the real marketplace to test the success and failures of a certain item. When the product does truly fail, the investment is lost and the item is quickly withdrawn (LaBarbera & Tucciarone, 1995). Researchers in many fields of marketing and advertising are looking to other methods of measuring emotion due to the fact that traditional research mainly understands rational thought, not true real-time responses (Harris, 2006). We hope to find equally effective and less expensive or risky results with physiological research. Self-report measures however should not be abandoned completely, however should be used to supplement and support new measures of consumer behavior (Liao et al., 2012; Picard & Daily, 2005; Sørensen, 2008). An application of behavioral recording devices will aid designers in the design process of packaging, prior to costly failures (Hurley, 2011; Tonkin & Duchowski, 2011; Tonkin et al., 2011).

Advanced (Physiological Methods)

Advanced measures of measuring emotion are outlined in order to move beyond self-reporting measures. Attitudes from consumers are known to be a poor predictor of in-store behavior. Instead, it is best to understand that behavior itself is the best predictor of success (Light, 1993). Daily and Picard look towards alternative methods of asking how people feel (Picard & Daily, 2005). Advertising industries are searching for aid in physiological devices to understand if their work is effective on consumers (Belch, Holgerson, Belch, & Koppman, 1982). Psychology, neuroscience, cognitive science, and neuroeconomics has documented the paths that emotions play an important role in the decision making process (Ahn & Picard, 2006). Today's implementation of physiological devices are outlined as affective systems to enhance the efficacy of learning and decision making (Ahn & Picard, 2006). Researchers in the fields of marketing agree that emotions should be investigated with physiologic devices, however an acceptable and realistic methodology must be found (Lee, Broderick, & Chamberlain, 2007; Stewart & Furse, 1982).

Through studying physiological means, designers are provided with more accurate methods of measuring consumer emotional response to designs when compared to self-report measures. Increased use of sensors and realistic experimental settings in the design process will help designers employ a transparent design process and strategy. Several different reactions can occur in a physical state when the body responds to certain perceived emotions including: change in heart rate, blood pressure, increase in sweat, respiration acceleration, and increased muscle tensions, facial expressions, textual expression, hand gestures, body movements (Consoli, 2009; LaBarbera & Tucciarone, 1995). We chose to focus on implementing EDA into the eye tracking evaluation workflow of packaging design due to its promising character traits, such as sensitivity to arousal, fast response times to stimuli, and mobile capabilities. Krugman states "physiological measurement isn't an exit interview. It's not dependent on what was remembered later on. It's a live blood, sweat and tears, moment-by-moment response, synchronous with the stimulus" (Krugman, 1981).

Bagozzi outlines three distinct advantages to real-time physiological procedures in 1991, including:

1. A close link of physiological measures to the mental and affective processes of a consumer
2. There is great potential in avoiding concealment or distortion of true consumer behavior.

3. An ability is created to monitor thoughts and feelings unaware to or hard to retrieve for the consumer (Bagozzi, Gopinath, & Nyer, 1999).

Physiological measures are recorded in real time during experiments instead of pausing and directly asking a question (Picard & Daily, 2005). To measure emotions in humans, researchers have relied on electroencephalography (EEG), heart rate variability, and electrodermal activity as a set of combined tools (Whang, Lim, & Boucsein, 2003). The results generated from physiological devices can be considered more valuable or truthful representations of consumer behavior to researchers because of the fact that consumers have little to no voluntary control over their autonomic nervous systems. An immediate and unbiased response in arousal can be quickly recorded through electrodermal activity systems (Bagozzi et al., 1999; LaBarbera & Tucciarone, 1995)

The importance of EDA recordings contrasted against self-report measures is highlighted with the research of Kohan in evaluating commercial stimuli. Kohan found that visible peaks occurred in the EDA data that closely followed the climax of commercials, however the verbal reports were found to not align in the same manner. Kohan's findings show potential in EDA research where the technique is able to bypasses the mind's "cognitive rationalizer," thus preventing false and biased consumer reports (Kohan, 1968).

Crowley's research builds a connection between physiological measures in packaging design where it is shown that it is possible for color to affect the certain perceptions of merchandise throughout a store (Crowley, 1993). Studies have shown that the color red is more physiologically and psychologically activating than other colors, like green. Investigations involving EEG, skin conductance, and self-report measures have revealed that red is significantly more arousing than green (Crowley, 1993). The following review will highlight the role of several human behavior based recording devices and how they connect to measuring emotions.

Eye Tracking

Eye tracking is a process by which the physical movement of the eye is recorded in order to investigate cognitive processes (Duchowski, 2007). Basic characteristics of eye movements are saccades and fixations. Saccades can be defined as the continuous movements of the eye. Fixations are the relatively still moments that occur in *still* points between saccades. Fixations usually last around 200-300ms. The term *still* is mentioned loosely in the context of the definition of the term fixation due to the fact that the eye is never really completely still. A constant tremor of the eye is present, termed nystagmus. Saccades can further be defined into three sub-groups: pursuit, vergence, and vestibular movements. Pursuit occurs as the eye follows a moving target. Vergence is when the eyes move inwards in order to focus on a close object. Vestibular movements occur when the eyes rotate to compensate for bodily movements. Fixations can also be directed into three sub-groups: nystagmus, drifts, and microsaccades. Nystagmus represents the previously mentioned constant tremor of the eye. Drifts and microsaccades are small eye movements slightly larger than nystagmus movements that allow the nervous system to compensate for the control of the human oculomotor system (Rayner, 1998).

The use of eye tracking in packaging design is now common, however additional devices and their acceptable methodologies have not been established in packaging design (Hurley et al., 2012; Hurley, 2011; Tonkin & Duchowski, 2011; Tonkin et al., 2011). Eye tracking has extended to consumer studies where it is found that an increased number of fixations on an object will enable higher recall of the object for the participant (Loftus, 1972; Tversky, 1974). In addition to understanding what consumers are looking at, and for how long, researchers have attempted to pair emotional attributes to eyes. A divide in eye tracking represents pupillometric measures as either emotional responses to stimuli or cognitive processing of information (Duchowski, 2007; Granholm & Steinhauer, 2004; Kuchinke, Võ, Hofmann, & Jacobs, 2007). Pupillometric measures, the changes of the pupillary aperture of the eye, have been utilized in eye tracking studies due to a theorized link to emotional responses to stimuli (Goldinger & Papesh, 2012; Granholm & Steinhauer, 2004). Belch supports the use of pupillary response as a measure of not only attentional responses in cognitive processing, but also a measure of affective magnitude, or arousal (Belch et al., 1982).

Bradley and colleagues conclude with a study that pupil response is directly linked to emotional arousal. This link was seen by measuring the change in pupil diameter through the presentation of arousing images (M. M. Bradley, Miccoli, Escrig, & Lang, 2008). It has been proven hard to replicate (M. M. Bradley et al., 2008), however Hess and Polt reported in 1960 that a person's pupils will constrict with the viewing of unpleasant images and dilate with the viewing of pleasing images (Hess & Polt, 1960).

A large difficulty that arises with pupillary research is that which concerns light. The primary function of the pupil is to respond to changes in illumination. Packaging design evaluations, especially those that are held in context shopping situations, make the variable of light difficult to control (M. M. Bradley et al., 2008). The stance of this thesis will explore eye tracking movements in relation to other physiological data. Pupillometric analysis will be disregarded due to the unclear reporting of past research and the difficulty involved with controlled lighting in the shopping context.

Electrodermal Activity (EDA)

Electrodermal activity can be understood as changes in skin conductance at the surface of the skin. As EDA measures arousal, this can be linked to emotion, cognition, and attention (M.-Z. Poh, Swenson, & Picard, 2010). It is documented that as early as 1888, scientists recognized a linkage between physical stimuli and the electrical activity of the skin (Andreassi, 2000). Hensel states that electrodermal activity (EDA) can be considered to be the most sensitive physiological indicator of events in consumer behavior (Hensel, 1970). EDA is the most frequently applied index of arousal in this area of research (Kroeber-riel, 1979).

Several terms are defined below that circulate within research surrounding electrodermal activity (EDA).

- EDA: electrodermal activity (Wolfram Boucsein, 1992; Dawson et al., 2007)
- EDR: electrodermal reaction (Groeppel-Klein, 2005)
- GSR: galvanic skin response (LaBarbera & Tucciarone, 1995)

- SR: skin resistance (Dawson et al., 2007)
- SC: skin conductance—the inverse of skin resistance (Dawson et al., 2007)

EDA is used as the main title of the process throughout this research. This is due to the fact that EDA is an updated term that reflects more than the others listed. EDA is a term that expresses the complexities of the skin as it transcends other terms like GSR in favor of a measure that goes beyond a "galvanic element" (Affectiva, 2013). Some words that can be used to describe arousal when measured by EDA can be understood as: activation, inner tension, alertness, excitation and energy mobilization (LaBarbera & Tucciarone, 1995).

Applications of EDA

Poh et. al provide a list of uses that a wearable EDA device can be applied towards including: psychopathy, dermatology, and neurology diagnostic purposes, depressive illnesses, prediction of functional outcome in schizophrenia, clearer distinction between healthy and psychotic patients, characterization of sympathetic arousal in autism, diagnosis of diabetic neuropathy, and also showing biofeedback to treat hyperhidrosis, epileptic, and psychogenic non-epileptic seizures (M.-Z. Poh, Swenson, et al., 2010; M.-Z. Poh, Loddenkemper, et al., 2010). EDA has also been used in sport psychology and physiology. For gymnasts, arousal regulation is important to enhance performances and to prevent injury (Knufinke, 2012)

Emotion Vs. Arousal

Two words that will be discusses frequently throughout the duration of this thesis will be arousal and valence. Arousal is a degree of activation (excited vs. relaxed) that can be experienced by a person. Valence is the descriptive term that is useful in understanding the degree of the arousal in terms of pleasantness or unpleasantness (positive vs. negative) (Whang et al., 2003). Human emotions, and the intersections of the tools involved in this research, can best be explained through Larson and Diener's "circumplex" diagram. It is limited in describing all emotions, however it is a good model in understanding the connection between arousal and emotion. This diagram (Figure 2) illustrates a two dimensional space based on the intersection of valence and arousal (Larsen & Diener, 1992; Whang et al., 2003).

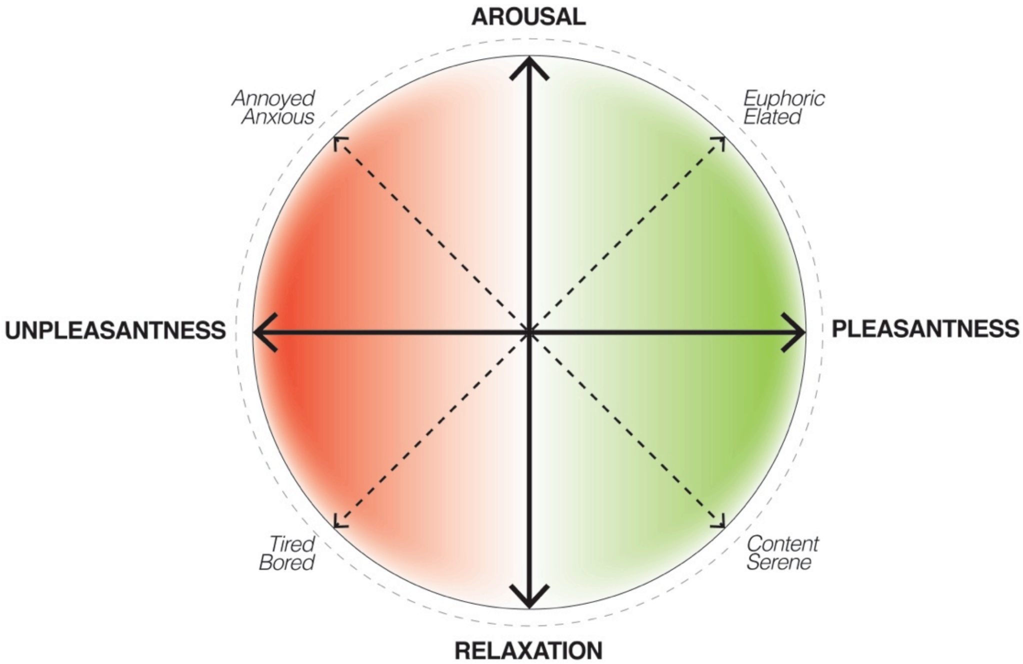


Figure 2: Two-Dimensional Model of Emotion [Adapted from (Larsen & Diener, 1992; Whang et al., 2003)]

Extensive tests have revealed positive results of the validity of EDA as a true measure of arousal (Kroeber-riel, 1979). Arousal is a basic part of human behavior (Kroeber-riel, 1979). In terms of psychophysiology, it is the basis of emotions, motivation, information processing, and behavior (Groeppel-Klein, 2005). Monitoring physiological process can produce definitions of emotional reactions and decision making (LaBarbera & Tucciarone, 1995).

Arousal is the first step to understanding emotion with the tool of electrodermal activity. Arousal can be understood as a global state of the mind. Arousal is a good indicator of brain activity. Arousal is not the same as emotion or items like stress and anxiety because it cannot specify the valence of a particular experience. Negative and positive characteristics must be paired with electrodermal activity measurements to define arousal as emotion. Other measures must be used to understand the valence of the arousal. Arousal can also be high in terms of high-negativity and, conversely, high-positivity (Knufinke, 2012).

Between participants, all physiological responses are universal. Similar and hard to mask, reactions occur when someone is aroused, all measurable through physiological devices. Direct measures of exterior motor skills in human behavior do not elicit the same, accurate responses across all participants. An example of the comparison can be understood when recording the activity of a person reading a book. Motor skills across participants can vary while physiological measures will accurately and sensitively detect changes in activation, or arousal (Kroeber-riel, 1979). EDA is a useful tool because it detects interest in items that people would not admit publicly. La Barbera summarizes this reasoning by providing the example that EDA is private (and more truthful), while a verbal response is public (being more influenced). EDA should be viewed as a useful aid to help consumers express their cognitive arousal (LaBarbera & Tucciarone, 1995).

EDA is considered to be valid with its high sensitivity, responsive to very small levels of variation (Groeppel-Klein, 2005). Emotions, specifically in the purchasing process, can be defined as a state of physiological arousal with added definitive cognitive aspects dependent on the viewed context (Consoli, 2009). Higher arousal leads to higher recall for a consumer (Kroeber-riel, 1979). Traditionally, research has revealed that EDA scores and consumer self-report measures contradict. It has also been found that EDA scores can be easily correlated to future marketplace performance. Therefore EDA can be considered a more accurate predictor of consumer behavior (Kohan, 1968; LaBarbera & Tucciarone, 1995). It is evident that arousal is an important aid in describing a consumer's buying behavior. Buyers are known to show higher arousal levels when compared to non-buyers (Groeppel-Klein, 2005). It is widely accepted across different disciplines that attention-grabbing stimuli and attention-demanding tasks are shown to evoke peaks in EDA responses (M.-Z. Poh, Swenson, et al., 2010).

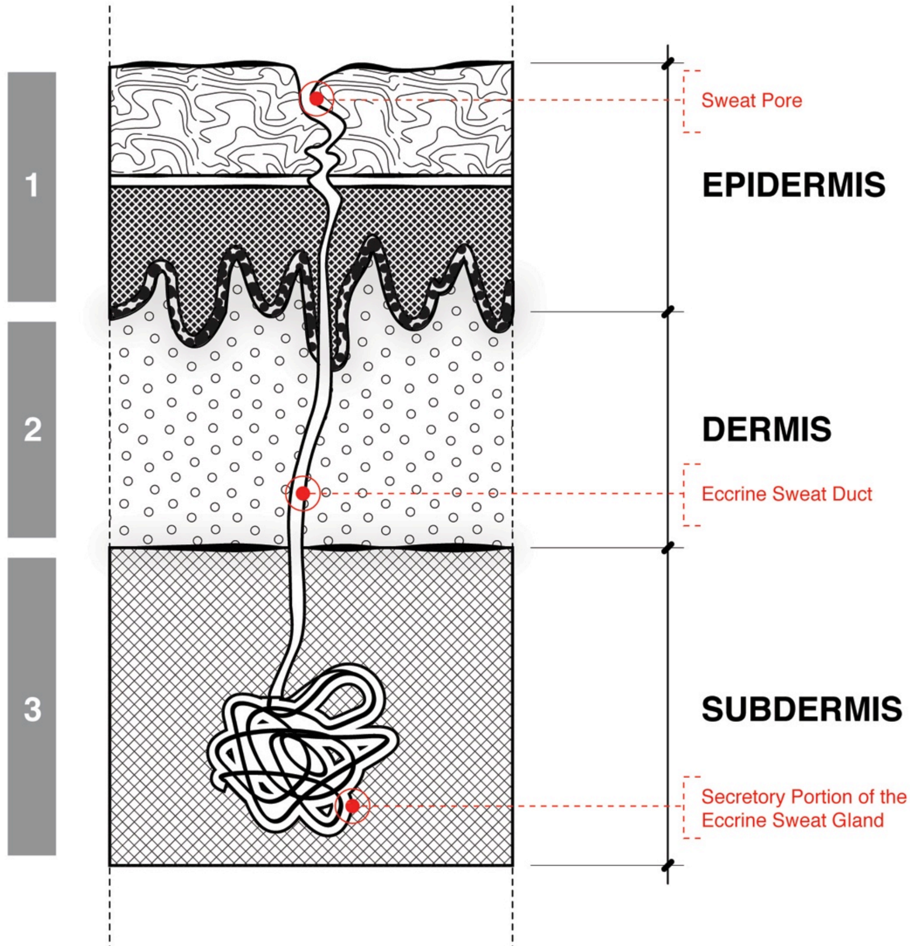


Figure 3: Anatomy of the Eccrine Sweat Gland through the Three Main Layers of the Skin [Adapted from (Dawson et al., 2007)]

A section of the skin can be viewed through Figure 3. The skin is constructed of two layers named the epidermis and the dermis, which both rest above the subcutis (or subdermis). (1) The epidermis lays at the surface of the skin consisting of epithelial tissue. This layer is more callous and rough as it becomes closer to the exterior surface.

This is the most important layer when researching EDA. (2) The dermis is much thicker, consisting of taut, fibrous connective tissue. (3) Inside the subcutis (or subdermis) is the secretory part of the sweat glands, fatty tissue, and supporting vessels. It is known that the epidermis becomes dryer as it becomes closer to the outside because the cells become less tightly arranged. There is however a permanent perspiration caused by the dermis through the epidermis. This permanent perspiration is even true when there is no sweat gland activity. This perspiration, or hydration, is useful in EDA because it depends not only on external factors, but it also depends on internal factors. This all leads to good electric conductivity of the skin, making two electrodes on the skin, valid methods of measurement (Wolfram Boucsein, 1992).

EDA recording devices work by applying a minor Direct-current (dc) to the outermost layer of the skin, beneath the electrodes that measure exosomatic measurements (M.-Z. Poh, Swenson, et al., 2010). There is no accepted standardization for recording sites in electrodermal activity. It is common for the electrodes to be placed on the palmar surfaces of the participant's hand. Usually one electrode is placed on both the middle of the index and middle finger of either hand. Sometimes it is found that researchers will place the electrodes on the area of muscle at the base of the thumb. Poh et. al's research is revolutionary in that they found that it is comparable to use the wrist as a measuring site for EDA. This allows the wearer of the technology to easily use their hands for manipulation within their environments. In addition to the electrodes interfering with the participant's daily activities, the traditional sites on the palmar region are shown to be highly susceptible to motion artifacts. By implementing measurements on the distal forearm, or wrist, one can more easily perform comfortable, long-term, and in situ recordings of EDA (M.-Z. Poh, Swenson, et al., 2010). Conveniently, the use of wrist as a measuring site for EDA has led to the applications of the wireless and unobtrusive Q-Sensor™ device created by Affectiva (M.-Z. Poh, Swenson, et al., 2010). The Q-Sensor™ is an applicable device to investigate the use of electrodermal activity in the shopping environment because it is a wearable technology suitable for movement and realistic conditions (M.-Z. Poh, Swenson, et al., 2010). Discrete monitoring is possible with the Q-Sensor™ due to its packaging on the consumer. The device is inconspicuous and non-stigmatizing because it is comfortable to wear and blends into one's culture as a simple watch or wrist bracelet (M.-Z. Poh, Swenson, et al., 2010).

Exosomatic recording is the application of electrodes to the exterior of the skin (the skin surface) to measure skin conductance (SC) in microSiemens (µS). Endosomatic is another method of measuring EDA, however this method is more invasive involving the placement of electrodes directly onto the sympathetic skin neurons (Groeppel-Klein, 2005). Device electrodes are usually placed on the wrist opposite to the participant's dominant hand. This can be assumed to allow the device to remain out of the way to reduce artifacts during experiments as participants write or perform other tasks with their dominant hand (Groeppel-Klein, 2005).

Arousal Illustrated

Two types of arousal can be identified including tonic and phasic arousal. Tonic arousal refers to long states of consciousness, which change slowly due to long events, or highly intensive encounters with stimuli. Phasic is more important for applications in packaging evaluation where this type of arousal responds to more specific stimuli, thus resulting in short variations in arousal levels. Phasic is beneficial in packaging evaluation because it shows the body's sensitive reactions, most closely related to attention (Groeppel-Klein, 2005).

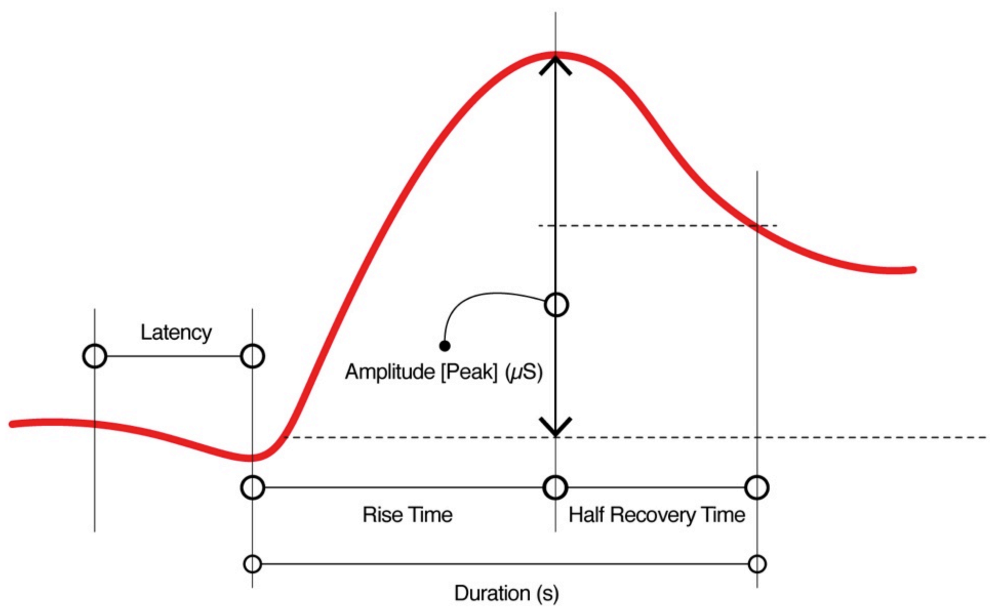


Figure 4: Graphic Diagram of Major EDA Response Components [Adapted from (Affectiva, 2013; Dawson et al., 2007)]

EDA can be easily measured wirelessly. Specific patterns in EDA data can be created, called artifacts, mostly caused by rapid movements, pressure on the electrodes, and change of distance between electrodes and the skin's surface. Normal walking however is not known to lead to artifacts in the EDA data output (Wolfram Boucsein, 1992; Dawson et al., 2007; Groeppel-Klein, 2005). An example of a peak in electrodermal activity (measured in microSiemens, μS) can be viewed through Figure 4.

EDA Limitations

There are limitations of EDA as a measure of emotion. First, EDA is capable of detecting changes in levels of arousal in a participant; however, the valence of the arousal is not distinguished. In other words, EDA alone cannot describe the positive/negative quality of an experience. Additional methods are needed to supplement the EDA recording such as other physiological measures or post experiment surveys and interviews (Groeppel-Klein, 2005; Picard & Daily, 2005). Next, response time in EDA is a complexity when comparing multiple participants and the data must be normalized. Researchers estimate that a person can experience a lag time in response to a stimulus of approximately 1-5 seconds (Latulipe, Carroll, & Lottridge, 2011). Finally, a wide distribution of responses in EDA data could be related to differences in sweat gland distribution across types or skin sudomotor innvervation (M.-Z. Poh, Swenson, et al., 2010).

Phasic electrodermal activity has proven to be a valid and appropriate indictor for measuring arousal, however many scientists claim that the technology has been not only widely used, but widely abused. Most studies including EDA urge for more research to be completed on the subject (Groeppel-Klein, 2005).

Supplemental (Neurophysiological Methods)

Researchers agree that there is a strong future for combining the physiological benefits of EDA with tools like fMRI (functional Magnetic Resonance Imaging), PET (Positron Emission Tomography), EEG (Electroencephalography), and EMG (Electromyography) (Davidson & Tomarken, 1989; Davidson, 1995; Nevid, 1984). With a future integration of neurophysiological tools we can locate and understand arousal directly from the brain. Neurophysiological measures can be understood as an objective replacement of traditionally collected, subjective, and declarative data. These tools must be combined in order to understand both conscious and subconscious reactions to stimuli (Ohme, Reykowska, Wiener, & Choromanska, 2009). With the implementation of the recording of brain activity, Davidson shows the promise of understanding valence with a non-self-reporting technique (Davidson, 1995). By using data provided from the brain, we could possibly eliminate the need for self-report measures and interviews all together to properly understand emotion (Groeppel-Klein, 2005).

EEG is the most versatile and sensitive measure available for detecting arousal in humans. Rhythmic fluctuations of the brain's electric potential are measured across a person's scalp. While this procedure is highly sensitive and revealing, it comes with limitations. Expensive laboratory equipment, complex data analysis, and stationary, motionless tasks make EEG a complicated and difficult measure (Kroeber-riel, 1979).

To clarify the valence of the reaction to the stimuli, electroencephalography (EEG), would be beneficial with EDA used as paired devices. Through studying left and right hemisphere dominance of the brain it is possible to add arousal degree and valence factors to the previously used EDA measures (Davidson & Tomarken, 1989). Consoli speaks about the ideas of a process termed "bio-feedback." A process by monitoring a consumer with neurophysiological measures (EEG) simultaneously with eye tracking to understand how a person felt while viewing a particular stimulus. No thoughts are applied towards the possible lag time that could be experienced with brain recording studies. Only a concept is provided (Consoli, 2009).

EEG is a useful tool in measuring a person's affective state. Davidson's research on EEG and emotions has revealed successful connections between the beta band of human brain EEG activity. The beta band (13Hz and above) serves as a measure of emotional valence. This activity has been termed the Frontal Asymmetry Paradigm where positive emotions have been shown to produce activity in the left frontal cortex of the brain, where negative emotions are known to activate the right side of the brain. Proper measures of EEG can be more successful than electrodermal activity measures. EDA requires additional post-study techniques to describe a particular valence (Davidson & Tomarken, 1989; Davidson, 1995; Ohme et al., 2009).

For EEG research, Davidson's model of emotion was used to locate the valence of reactions throughout the brain (Ohme et al., 2009). Ohme suggests that EEG be applied in parallel with the research methods of EMG and EDA. This is due to the reasoning that the degree of arousal measured from EDA can benefit from the direction of valence measured by EEG and EMG (Ohme et al., 2009).

The tool of fMRI (functional Magnetic Resonance Imaging) is known to be the most removed device from the shopping context. In fMRI studies consumers are removed from any close resemblance to a shopping context and placed inside a dark and hollow tube for observation (Reimann, Zaichkowsky, Neuhaus, Bender, & Weber, 2010). In a basic sense, fMRI measures the flow of blood throughout the brain, or “cerebral blood flow.” When a person reacts to a stimulus, a part of the brain is activated and blood is transported to the site of influence. This is exciting in that more emotional mapping can be discovered throughout the geography of the brain. A limitation with fMRI technologies is the lag time experienced from stimulus onset to brain activation due to the time needed for blood to travel in the brain (Vecchiato et al., 2011).

The importance of packaging and brand identity can be quickly understood with a study comparing Coca-Cola and Pepsi, which utilized fMRI technology (McClure et al., 2004; Morin, 2011). Most of the tested participants preferred Pepsi when they were blind-taste testing. When the same participants were told the brand of the soda, their preferences were immediately aligned with the brand Coca-Cola. Participant fMRI recordings revealed the activation of different regions of the brain through both preference decisions. This research example illustrates the strong connection that consumers have with packaging (Blakeslee, 2004).

Van der Lann et al. provide further research using fMRI where they use the data to aid consumers in making healthy purchase decisions. While previous research places brain research in a positive light, our paper provides alternative devices (eye tracking and EDA) to reduce research costs and increase the immersive shopping quality of the study participant (Van der Laan, De Ridder, Viergever, & Smeets, 2012).

EMG (facial electromyography) is used to evaluate the activity of the facial muscles (Ohme et al., 2009). EMG is a powerful instrument to measure the voluntary and involuntary facial expressions that might reflect both conscious and subconscious emotions (Ohme et al., 2009). Facial recognition can be used as a valuable and universal replacement for non-verbal communication where facial features are used to investigate arousal (Groeppel-Klein, 2005).

Electrodermal Activity Study Parameters

EDA holds great promise for being a reliable predictor of marketplace success (LaBarbera & Tucciarone, 1995). In prior years, when the development of EDA recording devices was immature, many researchers would report a high amount of "noise" throughout their data recordings. Noise can be considered as artifacts in the data produced from poor measuring devices, making analysis of the data near impossible. This lag in technology makes reason for its fairly new integration into packaging design fields (LaBarbera & Tucciarone, 1995).

For a package design to be effective, it is necessary for the item to evoke an emotional response from the consumer. This means that effective products will be known to create meaningful experiences in the consumer's personal life, an emotional area that is deeply rooted in the human subconscious. Once this emotional sector is triggered, the autonomic nervous system in involuntarily activated making EDA an applicable measuring device (LaBarbera & Tucciarone, 1995).

EDA, used as a measure of arousal in consumer studies, has been an interesting topic for more than 20 years, however the findings have revealed contradictory results, reason which could be due to many different methods used to measure the arousal. Historically, the validity of studies involving EDA have been questioned through more recent researchers (Groeppel-Klein, 2005; Rayner, 1998). The following chapter outlines typical conditions and parameters surrounding EDA studies and a hole is developed to represent a disconnect between desired and true behavioral results and falsified contextual representations.

Typical Study Environments

The question of this thesis research asks: How are we to collect realistic data from consumers, equal to that of actual consumer behavior, if the context of the studies is not parallel to that of real-life shopping environments? If we are to replicate the shopping context, is the use of electrodermal activity a valid measure?

Lang provides comments towards establishing a laboratory for measuring psychological responses to media messages. Traditionally, the main consideration of this type of space is that of control and solidarity. A proper laboratory is described as a controlled environment, one that is isolated from extraneous people, sound, and movement. Labs like this require computer monitors, televisions, and/or slide projectors to display experimental stimuli (A. Lang, 1994). It is understandable that studies should exist to isolate and control variables to increase validity, however people act differently depending on their surroundings (Tversky, 2003).

Due to EDA apparatus limitations in the past, studies have been limited to laboratory settings and artificial clinical environments (M.-Z. Poh, Swenson, et al., 2010). These measurements are restricted in short time segments inside the constraints of the labs and disconnected to the real world. In most cases the environment is not heavily criticized by the participant, however some researchers have documented the biases people have reported from participating in unreal, laboratory-like experimental surroundings (LaBarbera & Tucciarone, 1995). The advances that have been made in EDA measuring devices have opened opportunities of research towards applying the technology to the shopping environment. A wearable technology, such as the Q-Sensor™, is a suitable pair for consumer research towards packaging (M.-Z. Poh, Swenson, et al., 2010). The results gathered from a more mobile device allow a clearer picture to be drawn because the environment is not a limitation. Traditional, short EDA measures are not as reliable (M.-Z. Poh, Swenson, et al., 2010).

Groeppel-Klein concludes that the store atmosphere or environment should evoke reactions from consumers in order to attract their business (Groeppel-Klein, 2005). Opposite to the preceding claim, one experiment focused on the emotional affect of color. The floor and walls were carpeted with a standard gray color known to reflect all wavelengths of visible light. This was installed to allow no color effects to occur within the room itself. To help the participant believe they were in a realistic environment, the person was asked to imagine that they were in a furniture store (the environment of the tested stimuli). Unfortunately the participant was forced to remain static inside a fully unrealistic, imaginary environment (Crowley, 1993).

Another study, utilizing EEG and EDA together was conducted in an adapted room where all stimuli were presented on a television screen. Directly at the end of the study, participants were interviewed to collect any self-report data (Ohme et al., 2009). Similar studies employ large projection screens to display stimuli of various materials. Most always, participants are asked to remain stationary and are left isolated during experiments (Belch et al., 1982).

The most liberal of EDA studies found, to our knowledge, was that of Groeppel-Klein in 2005. This study involved an in-context grocery store produce test where participants shopped and EDA was continuously collected. The study however could have slimmed down with the usage of a less obtrusive EDA device. In addition to the large device held in a shoulder bag, the EDA was collected on the participant's palm, making it difficult to complete a normal shopping task (Groeppel-Klein, 2005).

Sample Sizes

Sample sizes in EDA studies are typically low due to the long time required to calibrate a participant to the device. A sample size of 30 is quite large in the physiological domain (Belch et al., 1982). One study utilized 45 participants (Ohme et al., 2009), while others were only able to draw conclusions from a pool of 12-15 (Groeppel-Klein, 2005). Additional research with a larger sample sizes would be worthwhile (Belch et al., 1982).

Participant Compatibility to EDA

Several factors can influence the measure of arousal including: personal introversion/extraversion levels, participant's interest in the product, sequence of exposure to stimuli, and the time of day (Kroeber-riel, 1979). Researchers have found that many individuals vary in responsiveness towards stimuli. With this range of recordings in mind, it is important for scores to be normalized across participants and not simply averaged evenly (LaBarbera & Tucciarone, 1995). Special care must be taken in the stages of data analysis to ensure even comparison across participants. In addition to personal character traits of introversion/extraversion, differences across participant EDA data could be due to differences in personal sweat gland distribution, also termed skin sudomotor innervation (M. Poh, Swenson, & Picard, 2009).

EDA Response Time to Stimuli

An interesting area of research with EDA is the reaction time measured from the actual onset of a particular stimulus to the recorded physiological measure. Several authors present typical response windows in EDA measurements. Dawson writes that a typical response window to stimuli can be regarded as 1-3 seconds after each stimulus onset (Dawson et al., 2007). LaBarbera & Tucciarone conclude that initial responses from stimuli have been recorded within the 0-2.5 seconds of stimulus presentation (LaBarbera & Tucciarone, 1995). . The largest window reported is from Latulipe and colleagues, showing a lag time experienced of 1-5 seconds (Latulipe et al., 2011). After measurements are collected, to clarify the source of the EDA response, researchers can employ interview techniques, retrospective playbacks, and additional self-report measures (Guan et al., 2006; M.-Z. Poh, Swenson, et al., 2010).

Determining Valence

EDA measures arousal, but not the valence, or perceived emotion derived from a personal response (like, dislike or positive, negative). Both very positive and very negative arousals can create equally large EDA responses (Ohme et al., 2009). A simple and effective solution to this is to conduct verbal post-study interviews with each of the participants. By conducting these interviews, more descriptive terms can be applied towards the valance of the arousal experience by the consumer (LaBarbera & Tucciarone, 1995).

To determine the valence of a degree of electrodermal activity, observations and questionnaires concerning mood and emotion can be used (Knufinke, 2012). While it is known that arousal plays a large role in consumer behavior and decision-making, it is difficult to indicate the valence of the measured arousal. Different self-report measures are used including PAD verbal scales, non-verbal color-pattern scales, and different interview methods. The color and pattern scale is interesting in that it removes verbal reasoning and allows the participant to choose a color or pattern that best represents an arousal level. However unique the testing procedure may be, this has not been used enough to support valid conclusions (Groeppel-Klein, 2005). Verbal scales were used in another study to confirm the consumer's valence. The experimental store was shown to cause higher arousal levels when compared to the control store. Verbal scales also significantly defined the term "joy" to be tagged with the layout of the store (Groeppel-Klein, 2005). A heavily used self-report measure to understand the emotions of consumers is the Self-Assessment Manikin (M. Bradley & Lang, 1994; Detenber, Simons, & Bennett, 1998; Liao et al., 2012; Morris, 1995).

Methods of EDA Calibration and Data Normalization

Response time in EDA is a complexity when comparing multiple participants. The data must be normalized. Important guidelines and statistical formulas must be used when considering skin conductance research (Ohme et al., 2009). In anticipation of normalization across participants, a uniform calibration process must be applied towards all participants. A successful calibration to an EDA device usually involves a set of exercises to create a proper moisture barrier between the skin and the electrodes. In addition to a proper moisture barrier, calibration techniques are required to activate the participant to a personal, maximum EDA level, used to normalize participants across groups (Dawson et al., 2007).

Researchers used the Q-Sensor™ in a study with gymnasts. One of their findings specified that the gymnasts were not affected by wearing the wrist devices (Q-Sensor™), instead of utilizing palmar sites. They also found that differences in the signal amplitudes could have been a result of the differences in eccrine sweat gland density across the participant's arms (Knufinke, 2012). In order to allow for maximum involvement in the environment, the EDA devices were applied to the participant’s non-dominant hand (Knufinke, 2012).

To "calibrate" participants, people can be asked "ice-breaking questions" to allow a sufficient amount of time to create a proper moisture barrier between the skin and the recording device's electrodes. Groeppel-Klein claims that this simple procedure created a baseline recording of zero for every participant. Not all people have the same skin conductance levels, therefore one can label Groeppel's calibrating techniques invalid and non-comparable (Groeppel-Klein, 2005).

Exercises on stationary bikes have been used to help their participants reach a maximum EDA level. This is completed as a quick 5-minute physical task. 10 minutes precede and follow the physical task to allow room for a baseline recording measurement (M.-Z. Poh, Swenson, et al., 2010). The process of blowing up a balloon has also been used to help participants reach a maximum recording (Dawson et al., 2007). One of the most aggressive and spotlighting techniques is accomplished where study proctors ask participants to loudly sing a portion of their favorite song (Chew, Mappus, & Jackson, 2010).

Several methods have been employed by researchers to normalize EDA data. This has produced a wide range of interpretations of EDA data across studies. Today's methods are much more conclusive than past methods. A formula for EDA data normalization is critical in the analysis across participants. One respondent's average EDA rating is not a valid indication of a final arousal score, a more accurate and acceptable form would be one that weights the scores across participants in relation to each person's individual skin conductance limits (LaBarbera & Tucciarone, 1995). It is difficult to find some researchers EDA data normalization formulas because of the need for companies to remain confidential (LaBarbera & Tucciarone, 1995).

Kroeber-Riel supports the use of the RELAMP relation to normalize EDA data between participants. The author claims that it tends to produce a normal distribution. The accuracy through stating this process of normal distribution appears low and on trial. To summarize, RELAMP (RELative AMPlitude) is a process of measuring the difference between the height of a peak (amplitude) and the EDA level prior to the peak. While it is useful to normalize the data, it is important to consider the maximum EDA values capable from the experiment as well.

An opposite method to normalize data can be followed here where, each response amplitude (peak) was divided by the participant's maximum recording. This time the minimum recording is not mentioned in this process of data normalization (W Boucsein & Hoffmann, 1979).

Additional common practices involve using the square root transformations to normalize response amplitude data (Dawson et al., 2007; Ohme, Reykowska, Wiener, & Choromanska, 2010). Lykken et al. propose a more in-depth solution to normalizing data across large groups termed, Range Correction. This involves computing the maximum range for a participant and comparing the peak as a proportion of his/her individualized range. The range correction procedure can reduce error variance and increase the power of statistical tests on the normalized data. The formula below represents the variables required of the Range Correction procedure to normalize data (Lykken & Rose, 1966).

$$\boldsymbol{EDA\ arousal\ ratio = \frac{(EDA_{peak} - EDA_{min})}{(EDA_{max} - EDA_{min})}}$$

$$\boldsymbol{EDA_{peak} = EDA\ recording\ at\ the\ peak's\ maximum}$$

$$\boldsymbol{EDA_{min} = overall\ personal\ minimum\ EDA\ recording}$$

$$\boldsymbol{EDA_{max} = overall\ personal\ maximum\ EDA\ recording}$$

Highlighted Case Studies

Chamberlain suggests that there is a "significant need to investigate the validation of current measures of emotions used in a marketing context" (Chamberlain & Broderick, 2007). No study, to our knowledge has been published with the use of mobile eye trackers and measures electrodermal activity in the shopping context. Several research agencies advertise the use of mobile eye tracking, EDA, and EGG, however their methodologies and findings are not available to the public (e.g. Neurofocus, Nielsen, Buyology). The following highlights of studies have helped support the path of this research, introducing mobile EDA and eye tracking to the consumer-shopping context. These are studies involving EDA, eye tracking or a combination of the two. A constant and critical question was asked during the construction of this review: what was the context of the experiment?

Hands-Free Gymnasts

Researchers are interested in using EDA in gymnastics to support theories that performers compete best at medium levels of arousal, compared to too high or too low levels. Injuries are also found repetitively at high levels of arousal. With more research of mobile EDA recordings in the performances of gymnasts, injuries can be decreased and, inversely, talent can be increased. This study qualified the use of the Q-Sensor™ EDA device to be used on the wrist of the participants. A hands free device enabled the gymnasts to compete more naturally. This same concept applies to packaging design evaluations in the shopping environment where it is best to allow the consumer complete freedom. A second major key point of this study introduces a concept to synchronize other events surrounding the EDA data. Gymnasts pressed the record button of the Q-Sensor™ in front of a camera to leave a marker for the synchronization of real-world events and the EDA data. This is applicable to the case of mobile eye tracking, where EDA data can be synchronized to participant's experiences (Knufinke, 2012).

EDA in Advertising

This study shows a promising connection between the data gathered from electrodermal responses and self-report survey data. Belch used a Pearson product correlation coefficient to determine if there was a relationship between the self-report and the physiological data. Results revealed weak correlations between the cognitive and physiological measures supporting LaBarbera's claim that physiological measures are usually more reliable predictor of marketplace success (Belch et al., 1982; LaBarbera &

Tucciarone, 1995). It is important to note Belch's use of a Pearson product correlation because this is a valid process to analyze and compare the two types of data.

Both cognitive and physiological measures were used to understand a person's emotional reaction to explicit imagery in advertising depicting nudity and suggestive human indecency. Three levels of nudity were displayed in the advertising stimuli. After viewing each stimulus, the participant rated each on a survey. Results indicated that these types of explicit images created affective responses with consumers measured with physiological devices (Belch et al., 1982).

EDA in Motion Picture

The emotional affect of motion pictures was analyzed on participants with the aid of self-report and physiological measures (EDA and heart rate). Moving and still versions of 27 different scenes were used as the emotional stimuli. The results concluded that motion pictures significantly increased arousal, supported by EDA and self-reported measures. The Self-Assessment Manikin (S.A.M.) was used as the self-reporting scale, used at the conclusion of each stimulus presentation. Participants also marked a line, where appropriate, to align closer to the words boring and interesting. The non-dominant hand was used as the site for the EDA electrodes (placed on the palmar site). The environment was not as crucial in this study, where participants rested in a comfortable armchair with stimulus presentation on a television (Detenber et al., 1998). A similar study is reported with the use of video news presentations in tabloid and standard formats (Grabe, Zhou, Lang, & Bolls, 2010). This relation to EDA and motion picture support

EDA's role in sensitively measuring real-time responses, an applicable trait to the freeform shopping context.

EDA and Color

Packaging designers are always drafting up color schemes of different varieties during the design process. This task is usually a personal and subjective decision made by skilled designers. Physiological measures could be useful in determining the colors for packaging design projects. Differences in color have been shown to produce changes in measures of heart rate, EDA, and self-report emotion surveys (Abbas, Kumar, & Mclachlan, 2005).

Crowley studies the physiological and psychological affects that various colors have in the everyday life of a consumer in shopping scenarios. Interactions between packaging, products, advertisements, and store environments are all bound together with the attentional effect of color (Crowley, 1993). Jacobs uses EDA, heart rate, and respiration rate to examine the arousing effects of primary colors. EDA was the only measure to report a significant finding, where red was listed as more arousing than blue or yellow and green more arousing than blue. 24 participants viewed the stimuli (colors) inside a normally lit room. The stimuli used in this study could have easily been presented as packaging stimuli. The participants viewed the primary colors while seated on a reclining chair. It would be interesting to see if these findings could be repeated with a freely moving participant inside the shopping context. The size of the EDA electrodes (0.75"-1.00") shows the age of this study at the year of 1974. EDA devices have largely progress since the 70s'. The next step of advancement should be that of the study context (Jacobs & Hustmyer, 1974).

EDA and Emotional Photographs

Images have a strong and affective power on people's emotional judgments (P. Lang, Greenwald, Bradley, & Hamm, 1993). A study involving colored photographs links closely to packaging design as positive images are always used to attract consumers through packaging. Photographs with varied valences from pleasant to unpleasant were used as stimuli to measure physiological and self-report responses. EDA, facial EMG, heart rate, and the Self-Assessment Manikin (S.A.M.) were used to aid researchers in examining the organization of emotional perception. 66 participants viewed the pictures for 6 seconds each while resting in a chair. As predicted, EDA increased with the ranked arousal from the self-reporting measures. A set of emotionally rated images was created to conduct further research named the International Affective Picture System (IAPS) (M. Bradley & Lang, 1994; P. Lang et al., 1993). This database of images should also be replicated on packaging as emotional stimuli inside the shopping context.

Neurophysiological Measures Reveal Subconscious Emotional Affect

A market research company looked at two different versions of a TV commercial for a skin care product. The two versions were very similar, however one included a very minute detail almost impossible to consciously notice. Neurophysiological measures of EEG (electroencephalography), EMG (electromyography), and EDA (electrodermal activity) were used to capture consumer reactions to the stimuli described. Post experimental interviews revealed that the difference was not consciously understood and reported. However, neurophysiological recordings revealed significant differences (Ohme et al., 2009).

A post experiment interview revealed significant differences in level of knowledge of the product with those that viewed the commercial ad with a suggestive hand gesture performed by a female model. This slight difference in the video supports the reasoning that strategic positioning can significantly enhance marketing communication. Ohme et. al approached neurophysiological tools for this experiment because of the fact that the difference is not consciously noticed. These scientists needed ways in which to measure subconscious reactions to stimuli. EEG was the main tool used this this study due to its very high temporal resolution (Ohme et al., 2009).

EDA was measured here using two 9mm diameter electrodes placed on the forefinger and middle finger of the participant's left hand. A third electrode, a reference point, was placed on the left forearm (Ohme et al., 2009).

The experiment was conducted in an adapted room for the study where all stimuli were presented on a television screen. Directly at the end of the study, participants were interviewed to collect any self-report data. Ohme and colleagues concluded that EEG, EMG, and EDA can measure different reactions to stimuli. This remains true even when the differences in stimuli are not consciously realized (Ohme et al., 2009).

Packaging Design with EDA and the S.A.M.

Liao et. al. completed a study using physiological measures to understand if consumers have unconscious emotional reactions to packaging designs. More simply stated, Liao researches whether packaging can evoke an emotion(s), which can be measured by physiological devices. There are two major differences between this study and this thesis. Liao did not employ the use of eye tracking and a realistic shopping environment was not used to test participants (Liao et al., 2012).

Liao provides a beginning foundation for physiological research in measuring emotional responses from packaging beyond eye tracking. Liao's study looks at reaction times in relation to package aesthetics (Liao et al., 2012). To understand if packages have the ability to evoke physiologically measured emotions, Liao looks at three elements of packaging design: images, colors, and typefaces. Liao employed the design topics into several different versions of a chocolate bar product. It is important to note that the brand name used was fictitious in order to eliminate any brand loyalties across participants. 120 participants viewed the stimuli while sitting down in front of a computer monitor screen. The computer was placed in a sterile environment. EDA was measured to record the degree of arousals, while facial EMG was employed to measure the valence of the emotional responses. The S.A.M. scale was used at the conclusion of the study to measure self-reported emotional valence towards the packaging stimuli (Liao et al., 2012).

To normalize EDA data, Liao calculated the EDA level by listing the difference between the maximum peak (1-6 seconds after stimulus presentation) and the minimum level (1 second prior to stimulus presentation). To slow the progression of the data, a square-root transformation was completed. While Liao is rather vague in the reporting of results, he summarizes that there was a significant effect on image and not on color or typeface (Liao et al., 2012).

Conclusions were made to develop the research of this thesis further: Liao found that physiological data can reveal emotional responses to packaging designs. Liao supports the future research of physiological measures in packaging. Thoughts of consumers traversing supermarket aisles in decision-making processes is briefly mentioned, however the testing of physiological measures inside the shopping context itself is not alluded (Liao et al., 2012).

EDA Measuring Consumer Environments

Groeppel-Klein concluded with several in-store experiments, that EDA is a practicable and valid way to measure arousal at the point-of-sale. The studies focused on asking the participant to perform a task (shop for an item) and then answer verbal questions to confirm their personal valence towards the reactions defined in the EDA data. This research is beneficial for starting a more detailed discussion on using EDA in shopping environments, however the research is not detailed enough to provide insight into the smaller scale of packaging. Many advocates agree that consumer testing must be completed in realistic and persuasive environments. We intend to dig deeper from the environmental and surroundings level, to the personal and intimate level of the package design. This paper supports the claim that EDA is a valid tool in evaluating consumer environments, however we would like to extend its causes to see if this applies to packaging. Can we pinpoint and measure the degree of a specific reaction towards a stimulus in a realistic environment (Groeppel-Klein, 2005)?

The Coupling of Physiological Measures and Design

By listening to the quiet conversations that consumers have with packaging, we can better understand how people make decisions affecting their health and well being. Eye tracking and electrodermal activity are two noninvasive and mobile devices that can document what draws attention, causing emotion to lead to action.

As the quality of these devices rapidly advance, it would be beneficial to add sensors into the real shopping environment to allow packages to globally adapt in real time to consumer reactions. A process of alteration could occur in real time in front of the shopper, highlighting and encouraging discourse between the consumer and the product. *Tesco's* virtual grocery shopping advertisements are close predictors of a future where consumers can shop in front of adaptive packaging (Simpson 2012). This example not only represents technology's place in the packaging design process but also in our society as a facilitator between humans and the products that enable us to live our lives.

Through measuring and understanding emotion in consumer purchasing decisions, designers can start to apply strategies to create an increasingly transparent design process. The methodology of this thesis is provided as a foundation for a new direction in the design process. Previous research combining packaging and physiological devices has not revealed a less opaque design process. Strategies similar to those in this thesis must be implemented such as: using a realistic retail environment, combining multiple measures (eye tracking and EDA), and data acquisition safety checks (post experimental video interviews and surveys). By using these realistic settings and interviewing the participant, interpreters of the data can confidently report the findings of the consumer studies. A future in physiological research will provide closer connections to consumers as evidenced through their purchasing decisions and responses. Many researchers are arguing for additional research to be conducted in applying GSR in studies, highlighting the fact that there is a dearth of academic work available (Groeppel-Klein, 2005; LaBarbera & Tucciarone, 1995; Liao et al., 2012).

CHAPTER THREE

EXPERIMENTAL APPARATUS

Two devices and one environment were used throughout this thesis research to measure the physiological responses of consumers through eye tracking and electrodermal activity. A pilot study (detailed in Chapter V) was first conducted in September of 2012 to introduce the Q-Sensor^{TM} device (manufactured by Affectiva^{TM}) into the packaging design evaluation workflow of CUshop^{TM}, *The Consumer Experience Laboratory*. Valuable lessons were gained from the pilot study that was applied towards a second study (detailed in Chapter IV) completed in February of 2013. This chapter of the thesis introduces and details the physiological recording devices and environment used in both the pilot and final study. Both studies are detailed in the following two chapters.

Tobii, Mobile Eye Tracking Glasses

Figure 5: Tobii Mobile Eye Tracking Glasses

Figure 6: Monocular Glasses

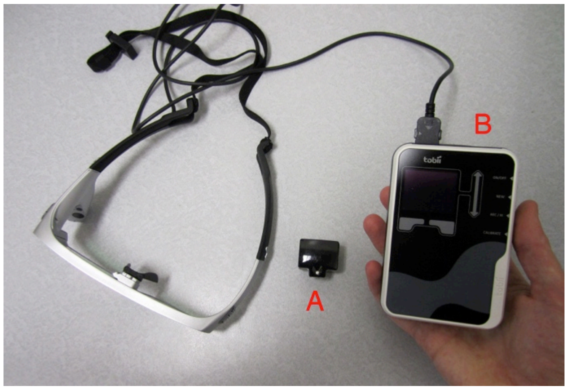


Figure 7: IR Marker (a), Recording Assistant (b)

Figure 8: Recording Assistant in Bag with Participant

The first recording device used in the EDA experiments was a mobile eye-tracker, the head-mounted *Tobii Glasses* (Figure 5- Figure 6). These glasses resemble ordinary glasses, however they are monocular, only sampling from the right eye. Two other pieces of hardware are used to record the eye movements: a handheld recording assistant and infrared markers (Figure 7). The recording assistant stores the data while the participant shops and also assists in the calibration process. The infrared markers are placed in a grid along the grocery-shelving units to delineate an area of analysis (AOA). Areas of Interest (AOIs) are drawn in *Tobii Studio* software, post experiment, to extract fixations for a specific area in the plane of the AOA (Figure 9). The sampling rate for the *Tobii* glasses is set at 30 Hz. In order to allow for freedom of movement throughout the shopping context, a shoulder strap and case was provided (Figure 8) to alleviate the need for the participant to hold the recording assistant, connected to the glasses with a small wire (Tobii, 2010). The eye tracking glasses are used in this study to locate the time of the first and last fixations of the participant on the packaging design stimuli.

Figure 9: Area of Interest (AOI) and Area of Analysis (AOA)

Affectiva Q-Sensor™, Electrodermal Activity

Figure 10: Electrodes on the Back of the Q-Sensor™

Figure 11: Affectiva Q-Sensor™ on Participant

Figure 12: Affectiva Q-Sensor™ in Shopping Context

An *Affectiva Q-Sensor™ Curve* was the unit used to measure electrodermal activity (Figure 10 - Figure 12). This is a mobile, lightweight, and wireless device that uses Bluetooth to stream real-time participant data. The sensor consists of a light box attached to the wrist of the user by an elastic wristband. One white button on the outside face of the device allows participants to time stamp events during the experiment. Two electrodes are located on the inner face of the device, which come in contact with the skin. The *Q-Sensor™* was set to sample at 32 Hz. The Q-Sensor™ was desirable to use in the packaging evaluation workflow due to the reasoning that it is wireless, unobtrusive, and fairly cost-effective at $2,000.00 USD. The Q-Sensor was viewed as a simple addition for consumers to wear as they walk normally in the shopping context. In addition to measuring EDA, the Q-Sensor™ measures acceleration and skin temperature. The white button (Figure 11) allows event marking in the data when processed post-study (Affectiva, 2013; Grifantini, 2011). By using the white, event-marking button, we were able to synchronize the EDA data with the video of the eye tracking recording. This synchronization helped to understand the EDA arousal levels in comparison to packaging stimuli throughout the shopping context.

CUshop^{TM}, The Consumer Experience Laboratory

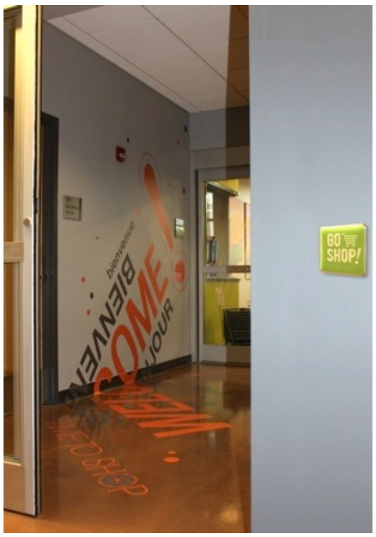


Figure 13: CUshop Entrance

Figure 14: CUshop Interior A

Figure 15: CUshop Interior B

Figure 16: CUshop Aerial View

Two studies were conducted in a realistic shopping environment, named CUshop^{TM} (Figure 13 - Figure 16), located on the Clemson University campus inside the

Harris A. Smith Sonoco Institute of Packaging Design and Graphics. The shop includes two double-sided shelving units, a frozen foods section, and an assorted produce section. The stimuli for both experiments were located on a full 16ft shelving section, all consisting of cereal breakfast products (Figure 44). Signs above each aisle directed the participants to the different products within the store (Hurley et al., 2012; Hurley, 2011; Tonkin & Duchowski, 2011; Tonkin et al., 2011).

CHAPTER FOUR

PILOT STUDY

Objective

The objective of the pilot study was to quickly assess the addition of the Q-Sensor™ EDA recording device into the packaging evaluation workflow. None of the results were found to be statistically significant, however trends were shown in the data that supported a further and more detailed investigation into the technology as applied to the shopping context. The following section outlines the methodology, results, and lessons learned from the pilot study. Where areas overlap, more detail can be found in the next chapter, the final EDA packaging study.

Hypotheses

Three hypotheses were constructed to pilot the use of EDA and eye-tracking as physiological tools in consumer shopping tests, used to gather responses to package designs inside the shopping context.

1. *Visible peaks in electrodermal activity will be produced when consumers experience emotionally designed packaging stimuli.*
2. *An increase in electrodermal activity can be related to higher purchase intent when the consumer terms the arousal as positive.*
3. *Eye tracking will reveal higher fixation counts per the negative packaging stimulus.*

Methodology

There were a total of 18 undergraduate participants (11 male and 7 female), ages 18-29 years (mean 21.9 years, median 22 years). Survey data confirmed that none of the participants suffered from glaucoma, cataracts, or any other eye impairments. Cereal boxes were chosen as the type of packaging stimuli due to the large surface area available for advertising graphics. Three different packaging stimuli were fabricated for the EDA experiment: Negative (Figure 17), Positive (Figure 18), and Neutral (Figure 19). The following keywords were used to design the positive package: *sustainable, all-natural, sunshine, happy, healthy, recycled, smiling, green, yellow,* and *blue*. The positive keywords were selected from popular claims found in current packaging trends. The positive package used graphic design elements that appeared to be happy and healthy while the negative package implemented contrasting graphics. Keywords including: *hazardous, warning, tetrachloroethane, harmful, dangerous,* and *black,* were used to influence the negative package's design. The negative keywords were selected as antonyms of the positive keywords. A 5-point Likert scale (extremely negative, negative, neutral, positive, extremely positive) was used in the survey to rate the design keywords. This self-report measure was taken to validate and describe the valence of any emotional responses recorded through EDA.

Figure 17: Negative Stimulus

Figure 18: Positive Stimulus

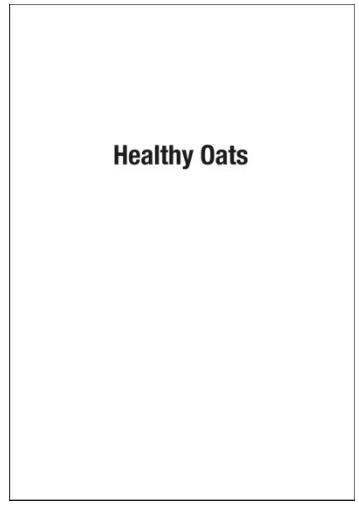


Figure 19: Neutral Stimulus

The participants were divided into four separate groups (Figure 21 - Figure 24). Each of the groups was exposed to one of the packaging stimuli (Negative, Positive, or Neutral). The first three groups experienced the packaging stimuli in an isolated cereal aisle where all of the surrounding boxes were white with simple brand name labels. The labels were all printed in the same size and color: black, Helvetica typeface (Figure 19 - Figure 20). The packaging stimuli were then placed in the direct center of the aisle at eye-level. The last group experienced something slightly different; the negative packaging stimuli were used within a modified context. Real, store-bought cereal packages replaced the white, surrounding boxes to create a more realistic cereal aisle (Figure 22). The negative package was chosen as the stimuli for this last group of participants because of its unnatural and highly contrasting design. If the negative package design would not be able to obtain an emotional response, recorded on the *Q-Sensor*^{TM}, then it would be hard to hypothesize a standard package emoting a response at all. A Witmer-Singer presence questionnaire was implemented in the post-experiment survey to understand if interaction within the environment was realistic to the participants. The questionnaire provided nine questions, arranged into the following sections (examples follow): involvement, immersion, sensory fidelity, and interface quality (Witmer & Singer, 1998).

- *Involvement:* *How completely were you able to actively survey or search the environment using vision?*
- *Immersion:* *Did you feel that CUshop was a realistic shopping experience?*
- *Sensory Fidelity:* *How much did your experiences in the environment seem consistent with your real world experiences?*
- *Interface Quality:* *How much did the control devices (wrist sensors and glasses) interfere with the performance of assigned tasks or with other activities* (Witmer & Singer, 1998)*?*

Figure 20: Positive Stimulus in Neutral Context

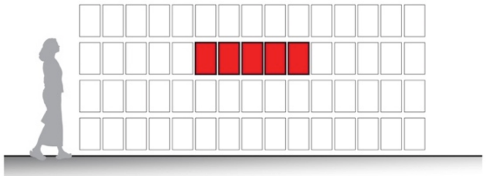


Figure 21: Negative Stimulus in Neutral Context

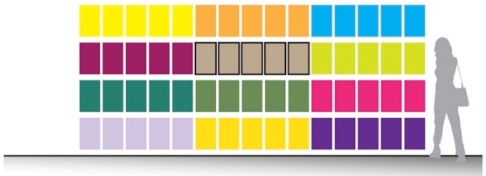


Figure 22: Negative Stimulus in Realistic Context

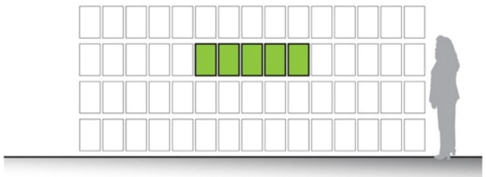


Figure 23: Positive Stimulus in Neutral Context

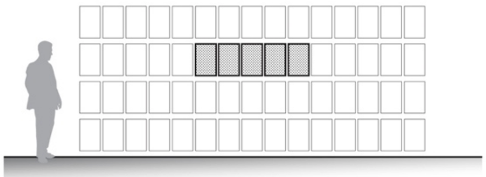


Figure 24: Neutral Stimulus in Neutral Context

Procedure

1—EDA Calibration:

The participant was first calibrated to the *Q-Sensor*^{TM} (Figure 11), needed to build an appropriate moisture barrier between the skin and the electrodes. Two sensors were used, one on each wrist. First, the participant was asked to sit in a sterile environment (a small conference room) and listen to 5 minutes of calming classical music. After 5 minutes, the experiment proctor would return and ask the participant to blow up a large balloon, *until it nearly pops*. This would force an increase in arousal consequently forcing the EDA reading on the sensor to spike. After the balloon exercise, the participant was left alone to listen to an additional 5 minutes of calming music. This allowed the subject to return to a baseline recording. The combination of relaxing music and blowing up a balloon would help establish a minimum (baseline) and maximum skin conductance level for each participant as well as allow time for moisture to build up on the participant's skin. An alternate method of EDA calibration is discussed in the following full-scale study.

2—Eye Tracking Calibration:

Following calibration of the *Q-Sensor*^{TM}, the participant was asked to stand 1 meter away from a wall, where he/she was calibrated to the Tobii eye-tracking glasses (Figure 5). The glasses calibration is a simple process where the participant wears the glasses and follows, with their eyes, the location of an infrared marker along the surface of the wall. Once the glasses were successfully calibrated, the participant was asked to press the *Q-Sensor*^{TM} button in front of their right eye. This was completed to log an instance on the eye tracking video and a marker on the *Q-Sensor*^{TM} data, needed for post experiment data synchronization. This concluded the calibration of the two devices to the participant.

3—Shopping Task:

Following successful calibration, the participant was provided a shopping list of 5 items on a clipboard and asked to "shop as you normally would shop." The five items were placed around the grocery store and appeared on the shopping list in the following order: (1) laundry detergent, (2) cookies, (3) cereal, (4) pasta, and (5) toilet paper. The subject would then be allowed to enter the CUshop^{TM} shopping environment to complete the experimental task.

4—Post Experiment Survey:

After completion of the shopping task, the participant was asked to remove the glasses and remain wearing the wrist sensors. A survey followed including four main segments: (1) basic demographic information, (2) a self-report extended response question, (3) a 5-point Likert scale on design elements (extremely negative-extremely positive), and (4) a modified *Witmer Singer Presence Questionnaire*. A self-report question (open response) started by prompting the participant to press their *Q-Sensor*^{TM} button. The participant was asked to press the button (EDA data marker) in order to see if there was a change in arousal experienced while answering the self-report question. An instance in the EDA data was marked, denoting the point at which the participant consciously reported their feelings towards the packaging stimuli. The question provided an image of the packaging stimuli previously viewed in CUshop™. The participant was asked to write in 2-3 sentences on how they felt emotionally towards the package. This survey question was paired with EDA to see if there was a change in arousal while consciously thinking about the stimulus. A 5-point Likert scale also asked them to rate the emotional quality of the design (extremely negative-extremely positive). The modified *Witmer-Singer Presence Questionnaire* was inserted into the survey to see how realistic the consumer felt when shopping around the cereal aisle, which at some times appeared uncommon with the selection of neutral white packages (Witmer & Singer, 1998). The full survey can be viewed in Appendix G.

5—Post Experiment Interview:

After completion of the survey, the participant was asked to participate in a 5-minute interview on their personal EDA data. The EDA data and eye-tracking video were quickly synchronized for immediate play back to the participant. Peaks on the EDA timeline were noted in the *Q-Sensor™* software as well as the *Tobii Studio* software. A video playback of the experience was shown to the consumer, starting one minute before and after each peak. The participant was asked to respond with personal thoughts in relation to each video clip. This interview data was used to confirm any emotional reactions to the packaging stimuli or the experimental design (Guan et al., 2006). The interview responses were compared to the survey results in order to confirm the emotional valence experienced. Most importantly, the interviews were conducted to see if the participant consciously noticed the stimuli presented.

Dependent Measures

Eye tracking metrics were studied to correlate the geographic location of the eyes to a participant's arousal levels. The eye tracking data was needed to confirm fixation on the package stimuli being studied. First fixation and last fixation on the packaging stimuli was recorded. Measures analyzed were *Time to First Fixation* (TFF), *Total Fixation Duration* (TFD), and *Fixation Count (FC)*. The electrodermal activity of the participants was used to illustrate the arousal levels of the participant throughout the study. Peaks (Skin Conductance Responses, SCR) were located using software provided by the manufacturer of the *Q Sensor*™ called, *Affectiva Q Analytics.* The software analyzed the EDA outputs (μS, microSiemens) of each participant, using peak detection to report the time, height (EDA), and duration of any abrupt increases in the skin's conductance (Figure 53). An EDA arousal ratio is calculated per stimuli (more detailed information can be viewed within the final study). The individual's EDA can be expressed as a proportion of their individualized range. This calculation, described as a formula below, normalizes the EDA data from person to person (Dawson et al., 2007; Latulipe et al., 2011).

$$EDA\ arousal\ ratio = \frac{(EDA_{peak} - EDA_{min})}{(EDA_{max} - EDA_{min})}$$

$$EDA_{peak} = EDA\ recording\ at\ the\ peak's\ maximum$$

$$EDA_{min} = overall\ personal\ minimum\ EDA\ recording$$

$$EDA_{max} = overall\ personal\ maximum\ EDA\ recording$$

Pilot Study Results

Due to the small sample size of the EDA study (n=18) the variances of the means were very large. Typically, studies involving EDA consist of 30 participants (LaBarbera & Tucciarone, 1995). The results will be presented in an exploratory nature for the pilot study. It is recommended that future studies involve larger sample sizes, as was completed in the final study.

Two self-report measures were used to evaluate the experimental design and to document the valence of the packaging stimuli. It is complicated to research EDA within consumer studies because of the difficulty found when isolating variables within a realistic environment. The experimental set-up is likely to suffer when variables are isolated, causing an unrealistic perception of the shopping scene. The results to the *Witmer-Singer Presence Questionnaire* report high involvement, immersion, sensory fidelity, and interface quality with no significant difference between the groups (Figure 25). It is interesting to note how participant presence decreases in most categories during the “neutral” stimulus, the most unrealistic retail construct. This trend of decreased presence with the “neutral” stimulus leads to the assumption that this group of participants was not fully immersed in the shopping experience. No experimental area appears to show significance; therefore we cannot report a decrease in reported presence, even though it may appear that way. Otherwise, as scores appear above the half-mark (2 out of 4), the conclusion can be made that the participants were generally immersed in the shopping environment. This helps in interpreting the validity of the eye tracking and EDA data.

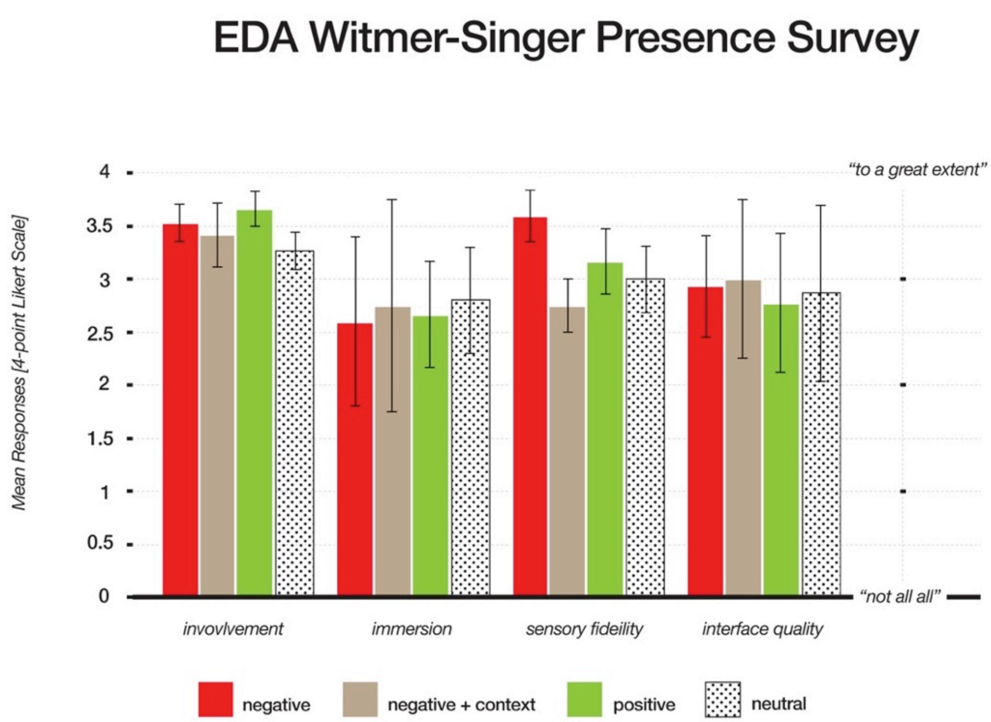


Figure 25: Witmer-Singer Presence Survey Results

The valence of the packaging stimuli was also evaluated (part 4 of the procedure), confirming the emotional associations towards the design keywords (Figure 26) and the packaging stimuli (Figure 27). Each of the three package designs is closely located near their respective area on the scale. The negative stimulus however was slightly below the normal, neutral line. The survey results conclude that the neutral package appeared to be more on the negative side of the spectrum. The negativity towards the neutral package is most likely due to the total lack of graphics, making the package feel abnormal and alien. The valence of 17 design keywords was surveyed to confirm the design decisions used to create the emotional packaging (Figure 26).

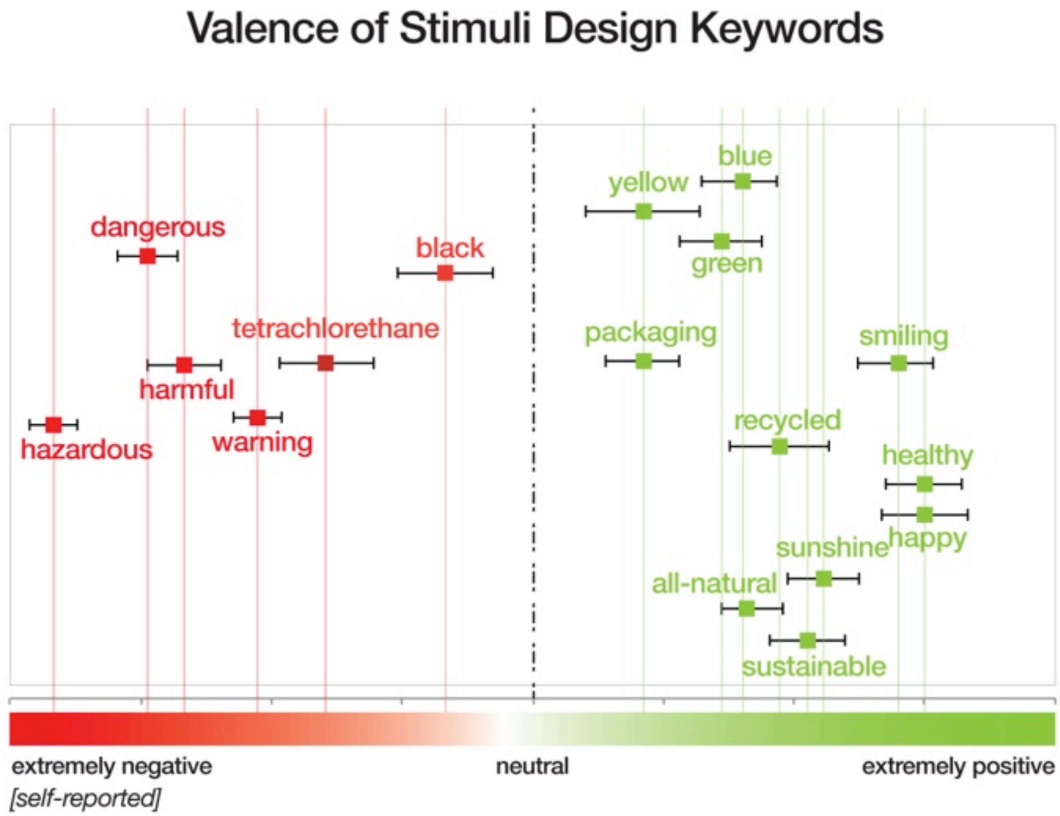


Figure 26: Reported Valence of Design Keywords

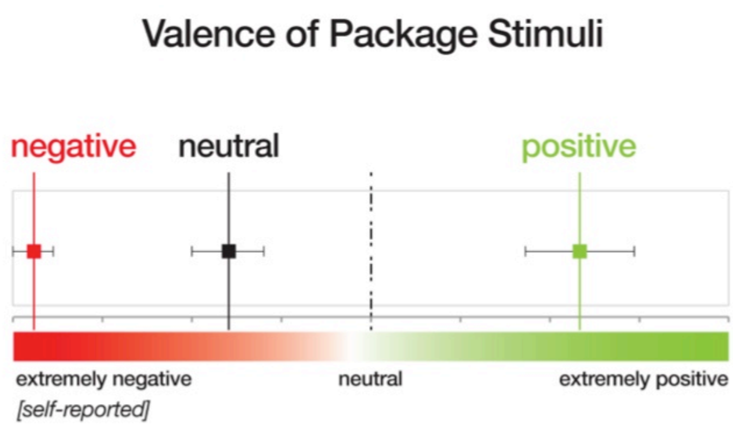


Figure 27: Reported Valence of Package Stimuli

Three eye-tracking metrics were analyzed with the packaging stimuli outlined as an AOI (Area of Interest): *Time to First Fixation, Total Fixation Duration, and Fixation Count*. The eye tracking metric most likely to provide significance would be the Fixation Count of the "negative + context" package stimulus (Figure 28*,* Figure 29). The "negative + context" package design, in this stimuli layout, did not loudly contrast with its surroundings as did the other stimuli layouts. An increase in fixations could suggest an increase in EDA or increase in mean peak height per stimulus. Total Fixation Duration (Figure 30), in addition to Fixation Count (Figure 28) represent the emotional power of the "negative" package design. At the critical time of the purchase decision, the negative package (both in neutral and realistic contexts) was the result of increased visual attention, confirming the third hypothesis. This could be attributed to confusion and unexpectedness towards the odd claims found on the "negative" design.

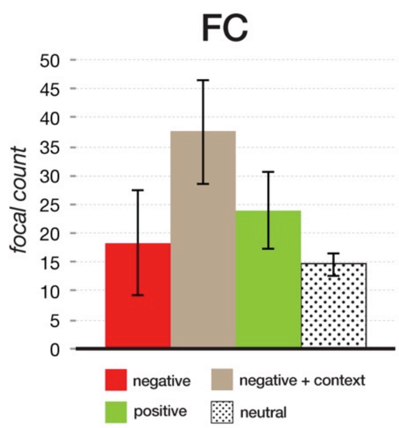


Figure 28: Fixation Count

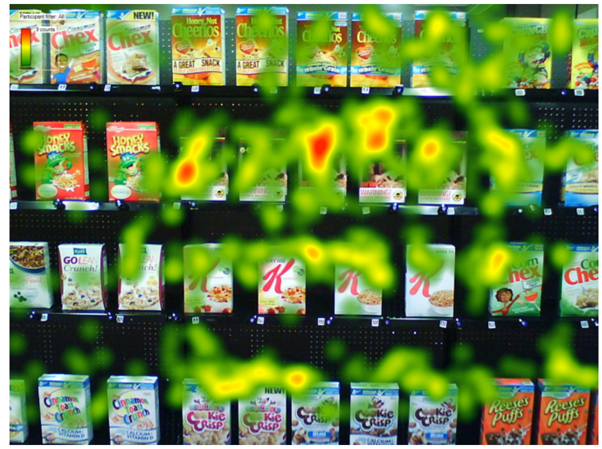


Figure 29: Aggregate Heat Map, Negative in Context

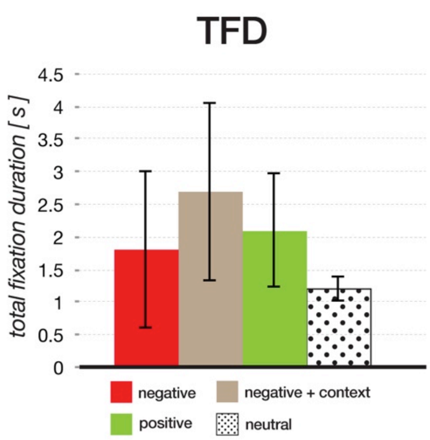


Figure 30: Total Fixation Duration

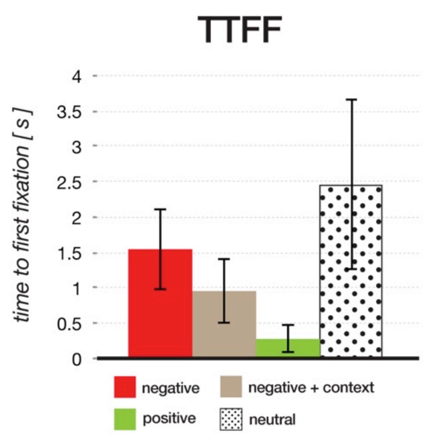


Figure 31: Time to First Fixation

The additional eye tracking metric of Time to First Fixation (Figure 31) represents the importance of positive packaging design. The "positive" package was the first design to grab the visual attention of the consumer. Conversely, the "neutral" design, the package with the least amount of design influence, was discovered last among the shelf.

The extreme increase in TTFF with the neutral package supports the importance of packaging design. Consumer shopping lists of the study provided interesting reactions where the "positive" design was the only package stimulus to be purchased (2 purchased, n=5). Aggregate heat maps (Figure 32 - Figure 34) illustrate the importance of emotional design influence in packaging. The red "hot" areas envelope the packaging stimuli, the only packages with designed graphics. Conclusions can already be made with eye tracking and self-reporting as to the importance of emotion in design at the moment of purchase. Eye tracking results show the strong importance of graphics displayed on packaging due to the shorter search times and increased viewing periods of different packaging designs. However, in order for the degree and accuracy of this emotional influence to be measured, EDA techniques should be implemented as an additional layer of trusted information.

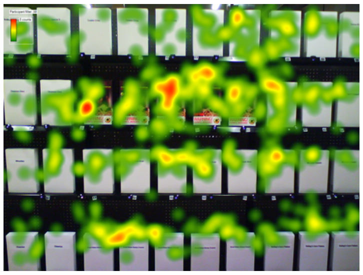


Figure 32: Negative Heat Map

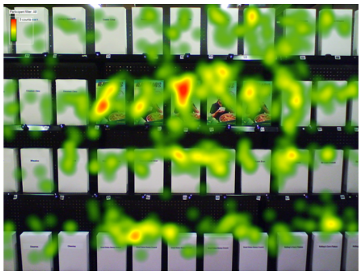


Figure 33: Positive Heat Map

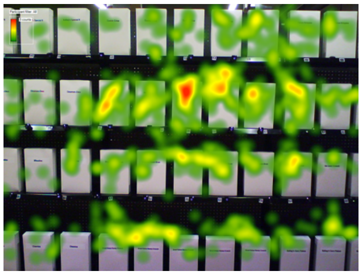


Figure 34: Neutral Heat Map

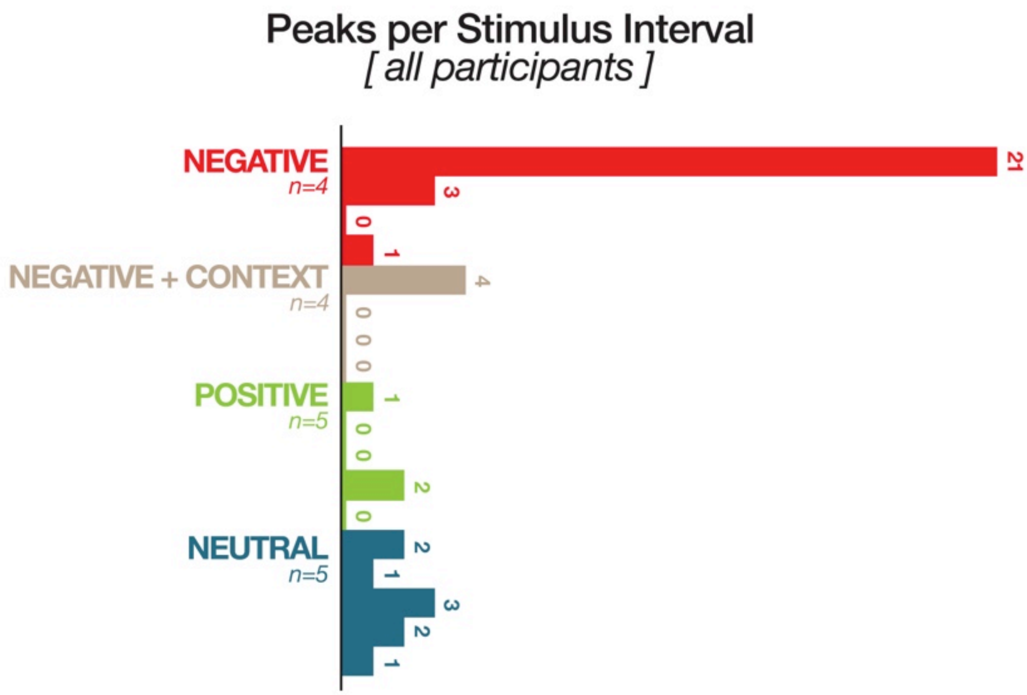


Figure 35: Peaks Per Stimulus Interval (experiences of each participant)

The above chart represents the peaks per stimulus interval experienced by the participants, all of which are represented by a horizontal bar (Figure 35). Some of the participants experienced as many as 21 peaks, while others experienced none at all. Through viewing the different experiences by each participant, it is evident that there is no obvious pattern developed in peaks per stimulus. The variances between the different stimulus groups are very large, showing an unpredictable pattern. The peaks per participant experienced during the neutral group is the most surprising, this is the only group where all of the participants experienced at least one peak. The neutral group is the only stimulus group highlighting a physiological response from each of the neutral stimulus participants. This could be evidence to suggest that the experimental layout of many white, neutral boxes caused a higher arousal than the positive and negative packaging stimuli. From this we can learn that experimental layout is highly important in recording EDA measurements.

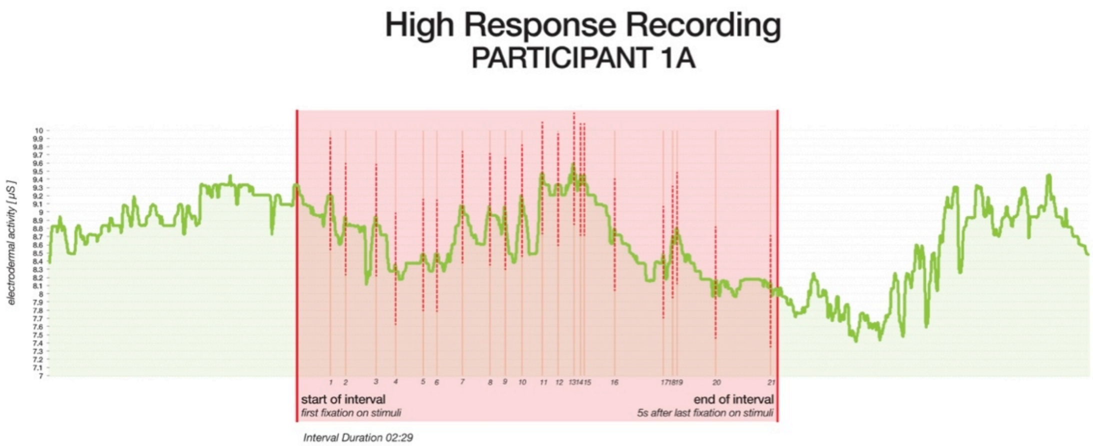


Figure 36: High Response EDA Example

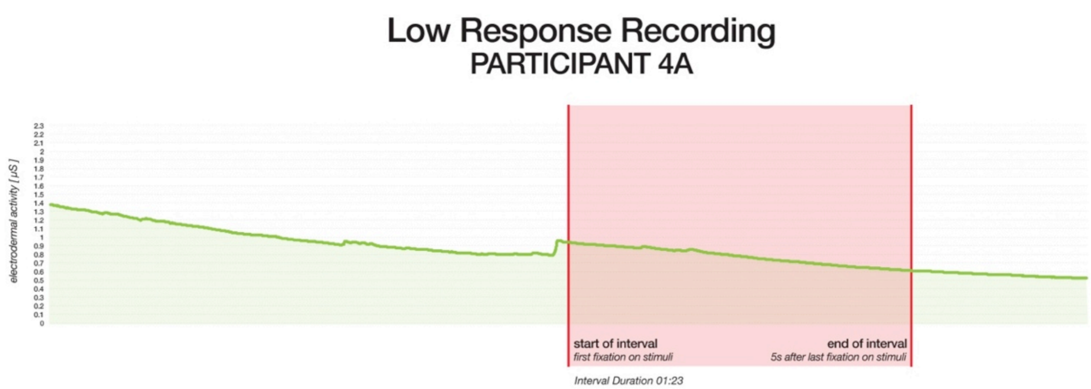


Figure 37: Low Response EDA Example

Figure 36 and Figure 37 represent the variability of the EDA recordings experienced across different participants (both participants experienced the same stimuli, the negative package design). Participant 1A is an illustrated example of a participant, whose EDA reading is highly active. Participant 1A experienced a total of 21 peaks, the largest out of the entire study. To an opposite affect, participant 4A showed a very low response in EDA. During the interval of stimulus viewing time, this participant experienced absolutely no peaks, and relatively few throughout the entire duration of the study. In addition to differences in body types, the variability of EDA responses could be attributable to the person's general arousal levels and personal character. Personal character by the definition of introversion and extroversion can also be linked to a person's EDA levels where some have very stable EDA (Bullock and Gilliland 1993, 113).

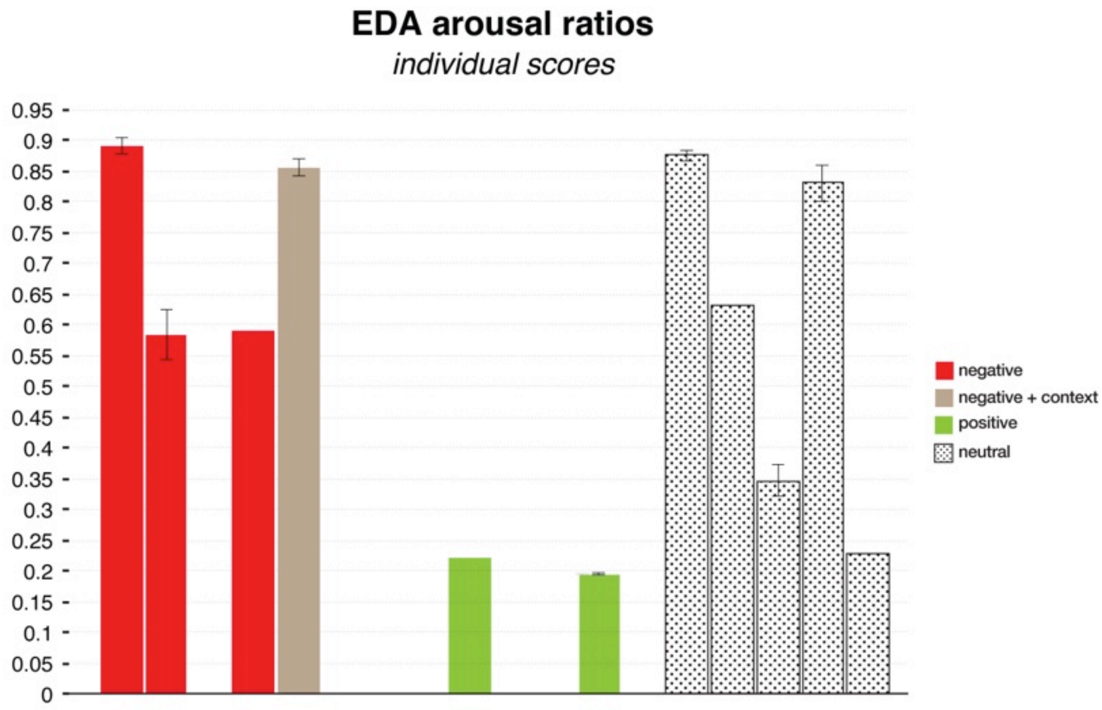


Figure 38: Mean EDA arousal ratios (experiences of each participant)

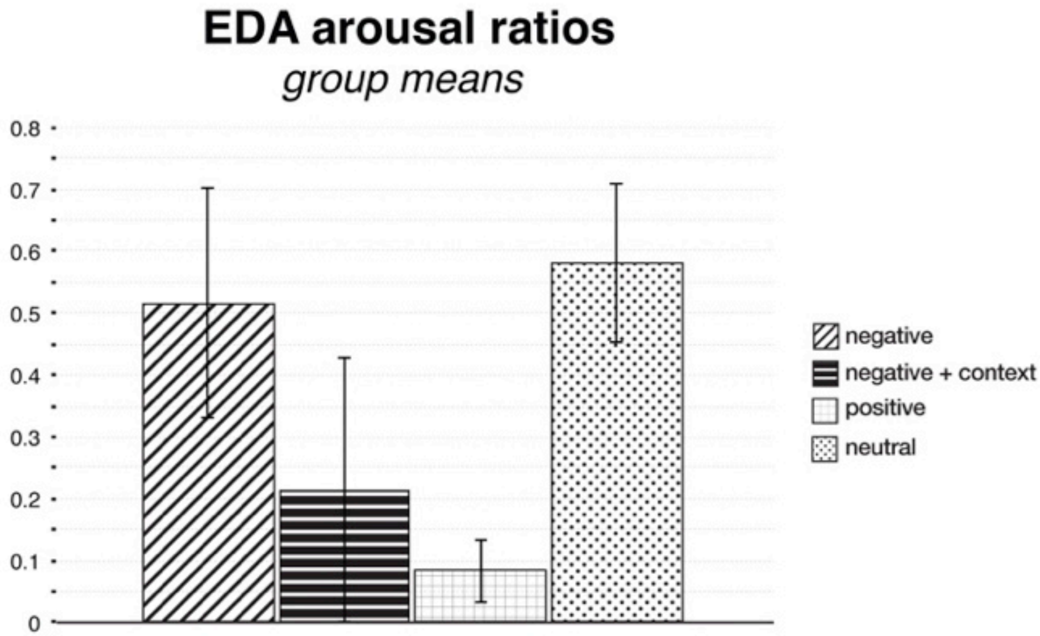


Figure 39: Mean EDA arousal ratios per stimulus

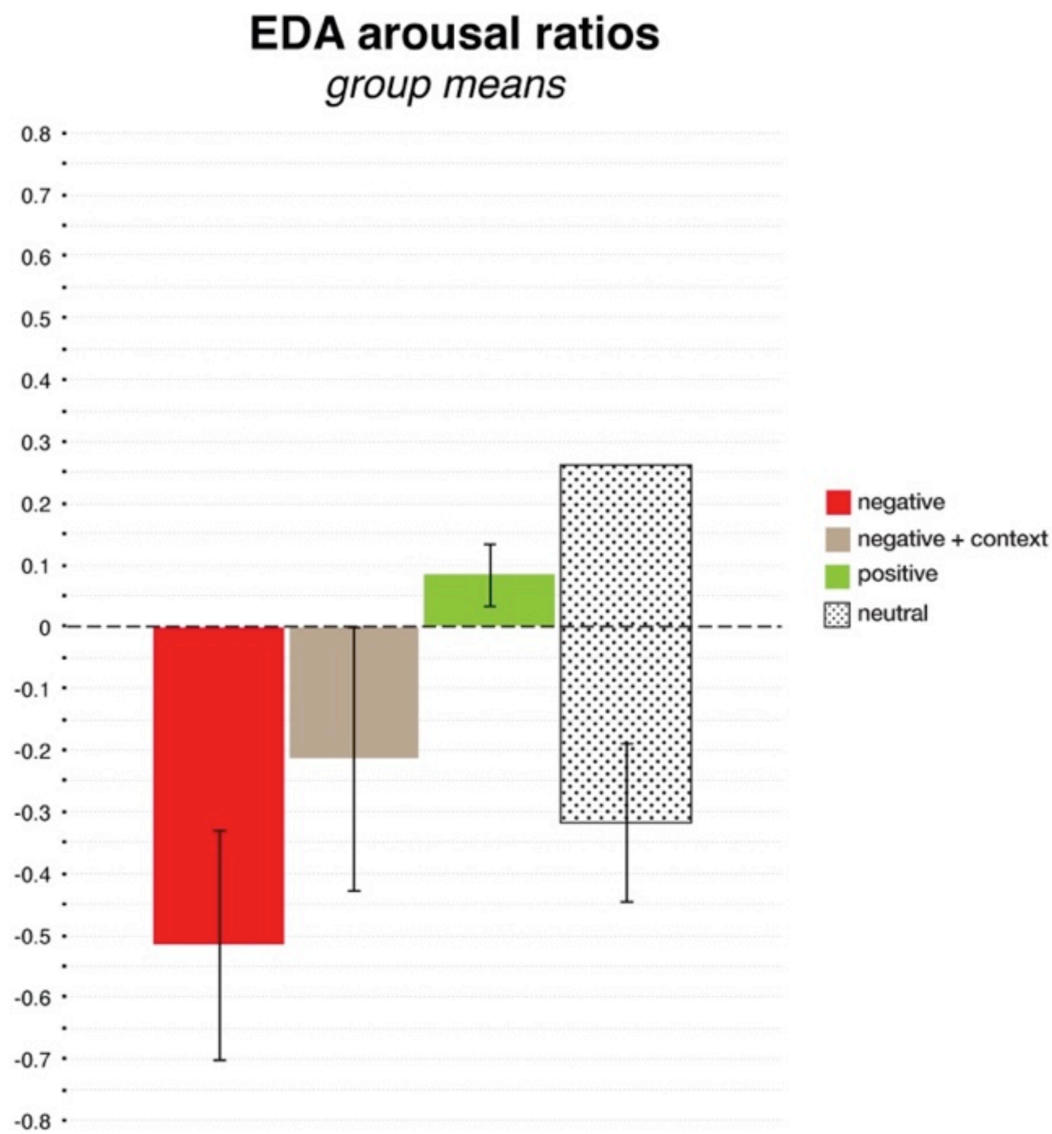


Figure 40: Mean EDA arousal ratios with valence

Careful calculations of the interval size for each participant was recorded to highlight the peaks experienced per interval. Some of the EDA peaks could have resulted from anticipatory or congratulatory emotions experienced by the participant before or after completing a task. This can be seen with almost each participant at the conclusion of each study where large peaks frame the end of the shopping task. Peak interpretation can sometimes be difficult and time intensive because post-experiment interviews and questionnaires are necessary to clarify the meaning behind peaks. The video provided from the eye tracking glasses is helpful in aiding participants recall the experience.

Figure 38 - Figure 40 are used to show the mean EDA arousal ratios calculated per stimulus. Figure 38 illustrates the individual EDA ratios per person, while also highlighting the variability of peaks experienced per person. Figure 39 summarizes the EDA ratios through the group means of each stimulus. In other words, Figure 39 shows the degree of arousal experienced per stimulus. To combine the efforts of Figure 27 and Figure 39, Figure 40 is drawn to show the degree and valence of the arousal experienced with each stimulus. With the combined data, this last graph represents the emotional reaction towards the contrasting packaging designs. As hypothesized (hypothesis 2) an increase EDA when the arousal is termed positive results in higher purchase intent. The negative package shows the largest increase in EDA, confirming the assumptions that a negative design will cause the largest arousal among the consumer—due to its strange and alarming uniqueness. The neutral design is found floating in the middle between negative and positive. As found through the EDA, eye tracking, and self-report measures, a successful design can be found in the positive *y* range of Figure 40. In other words, consumers are more likely to purchase positively designed packages (hypothesis 2) where negative packages are the most alarming and eye-catching (hypotheses 1 and 3).

Eye tracking metrics provide sufficient evidence to support the claim that package design plays a large role in emoting consumer reaction and increasing purchase intent. In eye tracking, the negative package was significant in increasing total fixation count over all packages due to its wildly different and abnormal design. This shows that consumers view designs differently. The neutral package has a significant difference in time to first fixation when compared against the positive package design. Positive designs (and designed packages alone) gain the consumers visual attention the quickest (TTFF). EDA and self-report measures represent the degree of arousal and high purchase intent of positive package designs. EDA and self-report measures also show high negative arousal to negative package designs with no purchase intents.

Pilot Study Conclusions

This study was successful in representing a useful connection between a mobile EDA device and the mobile eye tracking glasses for in context experiments. Eye tracking provides valuable eye movement metrics that explain the location and duration of a fixation. Electrodermal activity is a sensitive measure of arousal within a consumer. EDA is sensitive as in it is constantly measuring slight changes in a person's skin conductance at 32 Hz. Self-report measures such as interviews and/or surveys are required to describe the valence of EDA reactions to stimuli. The mobility of these devices, paired with a realistic retail setting enable researchers to collect definitive data.

Several lessons can be learned from this pilot study. The following is a list to build upon for improving the experimental design:

1. The design of the packaging stimuli could be considered arbitrary where positive and negative images were chosen from the personal view of the experimental

designer (M. Bradley & Lang, 1994; Chew et al., 2010; Detenber et al., 1998; P. Lang et al., 1993; P. J. Lang, Cuthbert, & Bradley, 1998).

2. The self-report surveys at the conclusion of the study were presented in the English written language. The expression of emotion in the self report measure would have been more successful had a visual method had been used to evaluate the visual stimuli. This is because the visual expression of emotion is a more universal language embedded with a clearer understanding across multiple participants. Written words, especially those that stand alone, can be understood in multiple ways depending on a person's background (M. Bradley & Lang, 1994; Morris, 1995).
3. The sample size (n=18) of the pilot study was considerably small for an appropriate EDA experiment. Usually sample sizes run close to 30 participants (Groeppel-Klein, 2005; LaBarbera & Tucciarone, 1995).
4. The balloon exercise performed in the EDA calibration segment of the procedure could be replaced with a higher arousing exercise. Such an exercise could be that of singing. The act of vocally singing is more strenuous and also typically increasingly worrisome for most participants when compared to using a balloon. By substituting singing as the calibration exercise, a more realistic range of EDA values can be collected per participant (Dawson et al., 2007).
5. The neutral package (Figure 19) used in the pilot study created confusion among most participants because of the lack of graphics and quantity of packages (Figure 20). In order to make the scene appear more realistic, we advise using realistic

cereal packaging with graphics and no neutral (or completely blank) designs. A realistic package is one that follows the graphic clarity and color variety of trending packages today.

6. The pilot study asked the participants to only choose one cereal. This is a weak part in the experimental design because there is little incentive for the consumer to spend a long duration along the aisle of the target package. When just one package was required in the shopping task, some participants would "purchase" the first cereal package to appear in their sight without viewing all of the options. Increasing the requirement to shop for more cereal can increase the chance for participants to see and experience the designed packaging stimuli.

CHAPTER FIVE

METHODOLOGY

Objectives

Is EDA, measured with the Q-Sensor^{TM} (a mobile and unobtrusive measure of arousal through EDA), a valid tool in recording arousal in consumers towards packaging designs when utilized in the realistic shopping context? Is the shopping context too "noisy" to adequately interpret EDA data received? What other measures can be used to accurately measure and report emotion in packaging studies? To summarize the link between physiological data and self-report data, a three-dimensional model is proposed to help aid in the visualization of the complex emotional results investigated within this thesis.

Hypothesis

The Null Hypothesis, H_O:	Electrodermal Activity is not valid for measuring arousal to packaging stimuli in the shopping context.
The Alternative Hypothesis, H_A:	Electrodermal Activity is a valid tool in measuring consumer reactions to packaging stimuli when situated in a realistic shopping context.

Dependent Measures

To help answer the hypothesis above, the following points summarize the dependent measures: (1) EDA arousal ratios per the positive and negative package designs, (2) Total Fixation Duration on both stimuli, (3) Time to First Fixation on both stimuli, (4) correlations between the EDA ratios and S.A.M. self report measures, (5) correlations between EDA ratios and eye tracking data, and (6) differences between EDA levels of introverts and extroverts.

Stimulus Design

Figure 41: Death Crunch, Negative Stimulus

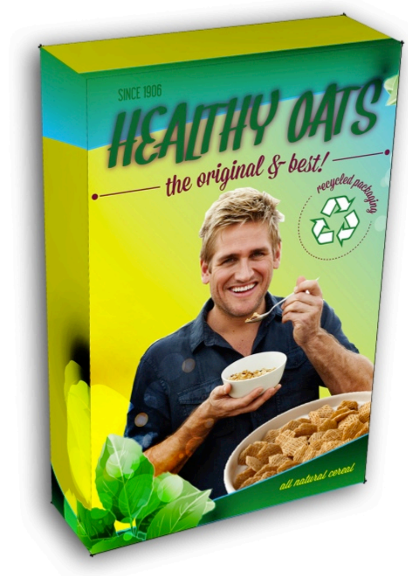


Figure 42: Healthy Oats, Positive Stimulus

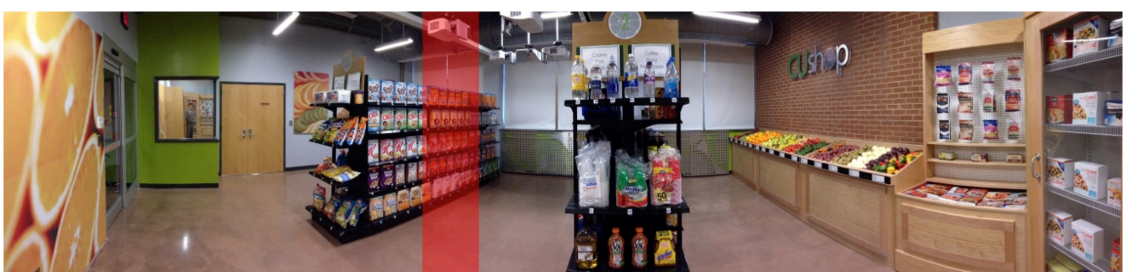


Figure 43: Stimulus Location Highlighted in Red

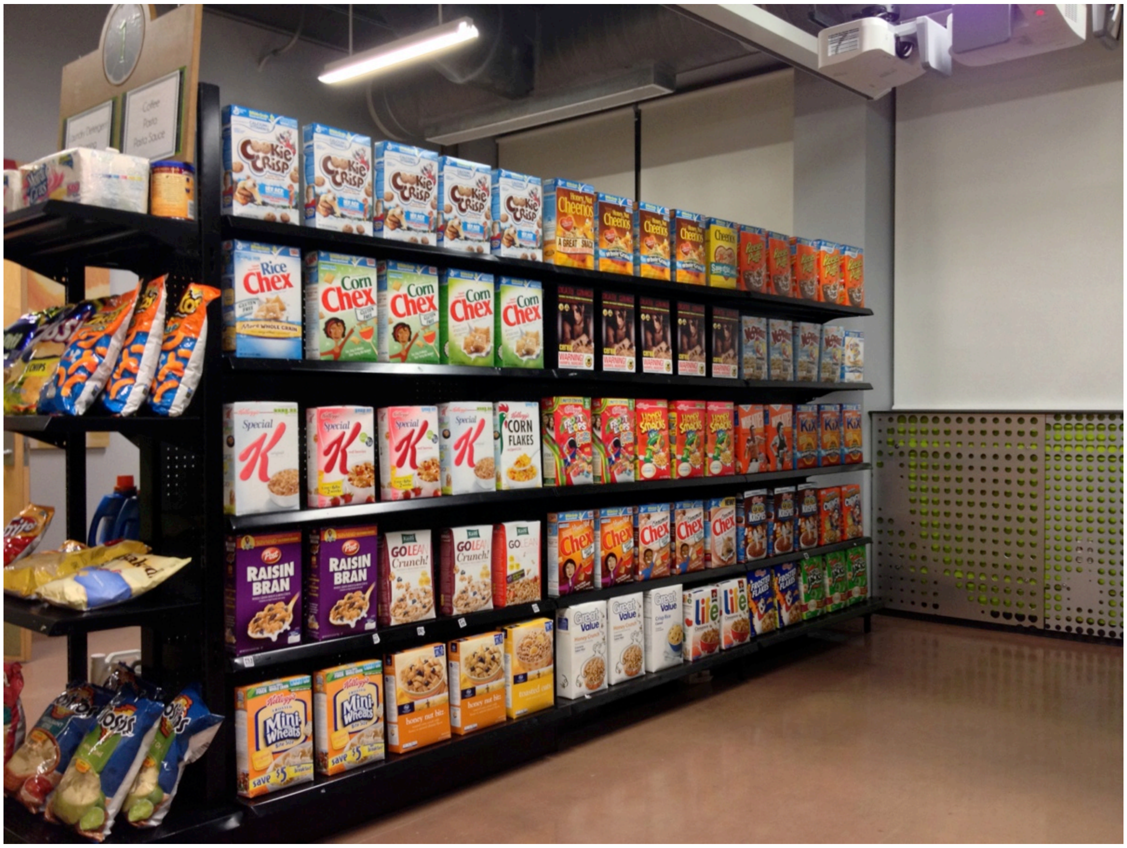


Figure 44: Negative Stimulus in Shopping Context

Cereal boxes were chosen as the type of packaging stimuli due to the large surface area available for advertising graphics. Two packaging stimuli were fabricated for the EDA experiment: Negative (Figure 41) and Positive (Figure 42). The following keywords were used to design the positive package: *sustainable, all-natural, sunshine, happy, healthy, recycled, smiling, green, yellow,* and *blue*. The positive package used graphic design elements that appeared to be happy and healthy while the negative package implemented contrasting graphics. Keywords including: *hazardous, warning, tetrachloroethane, harmful, dangerous,* and *black,* were used to influence the negative package's design. A 5-point Likert scale (extremely negative, negative, neutral, positive, extremely positive) was used in a survey to rate the design keywords. This self-report measure was taken to validate and describe the valence of any emotional responses recorded through EDA.

An online survey, utilizing *Survey Monkey*, collected 126 responses from people on their emotional reactions to the packaging stimuli (Appendix I). Survey Monkey recruited the 126 participants. Each participant that completed the survey was rewarded with $3.00 USD as an incentive. The emotions were recorded with the S.A.M.: *Self Assessment Manikin Scale* to determine the emotional feelings associated with the two packaging stimuli. The survey included one random package design at the beginning that was used to accustom the respondent to the S.A.M. scale. There are three scales in the S.A.M. including happiness, arousal, and dominance. The following diagrams (Figure 45) represent the S.A.M. characters. To complete the survey, the online participants selected the manikin that best represented their feeling towards the packaging design pictured above the scale. The same S.A.M. survey questions were completed by all eye tracking and EDA study participants at the conclusion of the shopping task.

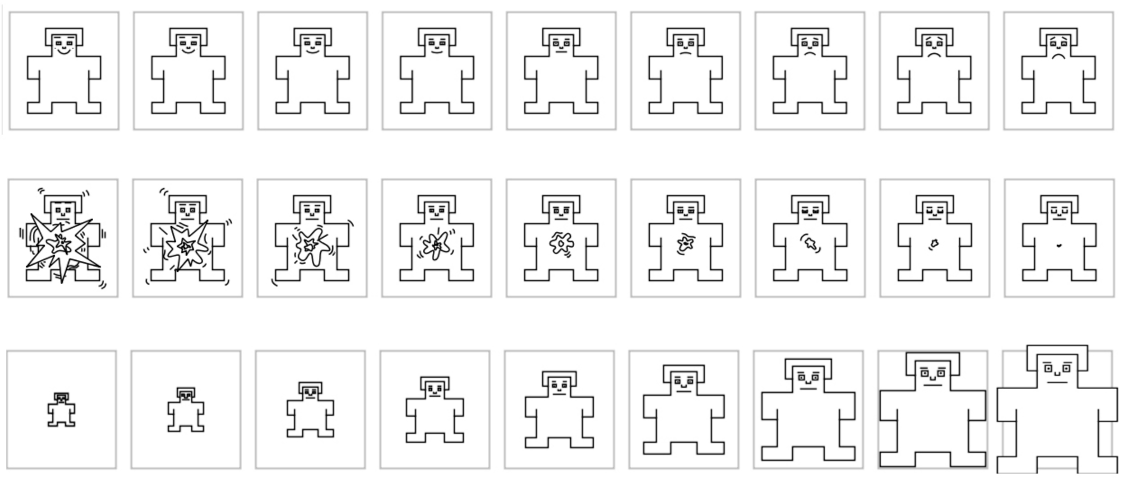


Figure 45: Self-Assessment Manikin Scale (S.A.M.). From Bradley and Lang, 1994, © Elsevier Science Ltd., UK, used with permission.

Stimulus Fabrication

Figure 46: Roland LEJ-640 (printing)

Figure 47: Kongsberg iXL44 (crease and cut)

The packaging dieline of the cereal box stimuli was created in *Esko Artios CAD 2012*. Digital files of the packaging graphics were created using *Adobe Photoshop CS5.5* and *Adobe Illustrator CS5.5*. The packaging stimuli were physically created in the prototyping lab located at the *Harris A. Smith Sonoco Institute of Packaging Design and Graphics* at Clemson University. Here, the packages were printed on 12pt paperboard on a *Roland LEJ-640 UV* digital inkjet printer (Figure 46). After printing on the flat-stock material, the printed sheet was creased and cut on the *Kongsberg iXL44* cutting table (Figure 47). Once all the packages were creased and cut, they were finally assembled with hot glue adhesive.

Stimulus Layout in Store

Figure 48: Stimulus Layout in Context (dotted X's represent negative and positive stimuli placement)

The participants were divided into two separate groups of 21 participants each. Each of the groups was exposed to one of the packaging stimuli (Negative or Positive). These groups experienced the packaging stimuli in a cereal aisle where all of the surrounding boxes were normal and familiar cereal packages. The packaging stimuli were then placed in the direct center of the aisle at eye-level (Figure 48). Each cereal type was grouped in duplicates of five each. To effectively contrast from the pilot study, only realistic cereal boxes were used to make the shopping context extremely lifelike and current.

Procedure

Sample Size

There were a total of 42 undergraduate participants (26 male and 16 female), ages 18-29 years (69.05% = 21-29 years old, 30.95% = 18-20 years old). The participants were undergraduate students receiving extra credit for their participation. Survey data confirmed that none of the participants suffered from glaucoma, cataracts, or any other eye impairments. 42 participants were used in this study because it was desired to have at least 30 total participants. 12 people were added to the schedule to make up for any bad data collected either through faulty calibration or unusual circumstances.

EDA Calibration

To calibrate to the Q-Sensor™ an appropriate moisture barrier needed to be established between the skin and the electrodes. Participants first placed the EDA device on the wrist opposite to their writing hand so as to reduce the chance of bumping or moving the sensor during the shopping task (Ohme et al., 2009). After securing the device, the participant was left alone inside a blank and sterile room to listen to five minutes of calming, classical music. After the five minutes, the study proctor returned to the room and asked the participant a somewhat awkward question. The participant was asked to sing their favorite song. The script was performed as follows:

> *Please think of one of your favorite songs. In 20 seconds I need you to sing the chorus or main verse of the song as loudly and clearly as possible. You may start when you are ready.*

The purpose of the singing exercise was to find a maximum arousal level, easily caused by high stress and anxiety (Dawson et al., 2007). Next, after the singing exercise, the participant was left alone again to listen to three minutes of the same calming, classical music. An example of the rise and fall created during the EDA calibration session can be viewed in the blue shaded area inside Figure 49. This concludes the calibration of the participant to the EDA *Q-Sensor*^{TM}.

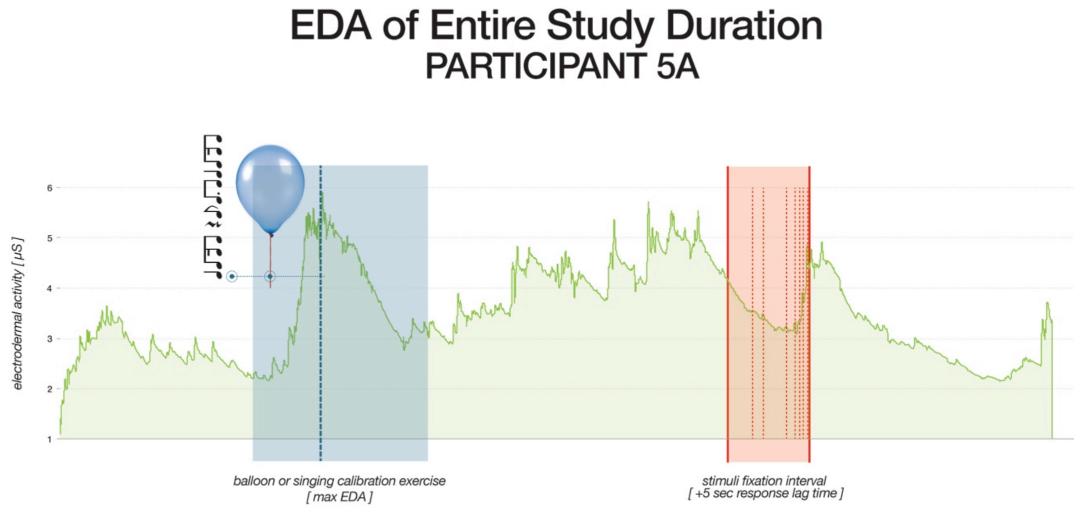


Figure 49: Example of EDA peak experienced during calibration

Eye Tracking Calibration

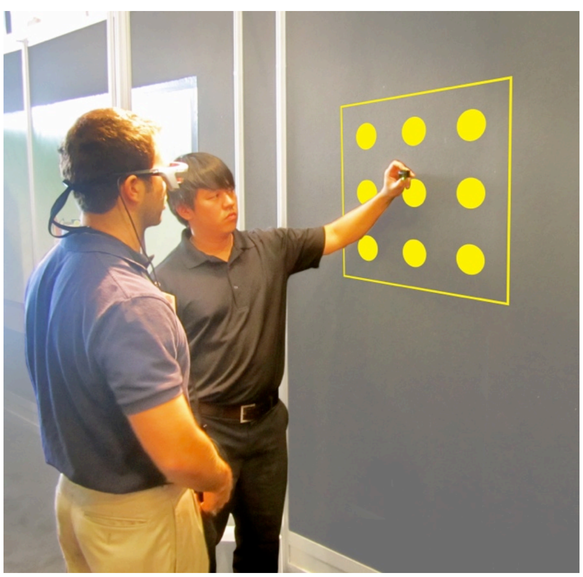


Figure 50: Eye Tracking Calibration

Figure 51: Q-Sensor^{TM} Time Event Marking

The calibration of the eye tracking glasses is a very quick process. The participant is asked to stand one meter away from the wall. The study proctor follows automatic prompts on the Tobii recording assistant to lead the participant through the calibration process. To complete the process, the participant is asked to remain still while only moving their eyes. The proctor moves a small IR marker along the wall in the illustrated pattern (Figure 50). As the proctor moves the IR marker, the participant follows the motion with their eyes. After following the nine imaginary dots, the participant is calibrated. Before sending the participant into the shopping environment, one last step is crucial in the calibration process. The participant is asked to raise their EDA device in front of their eye tracking glasses. Once the wrist device is in front of the eye tracking camera lens, the participant is asked to lightly press the button on the *Q-Sensor*™ (Figure 51). Once the button is pressed, a light shines and can easily be found on the eye tracking video. This critical step is the method of synchronizing the two streams of data.

Shopping Task in Physical Context

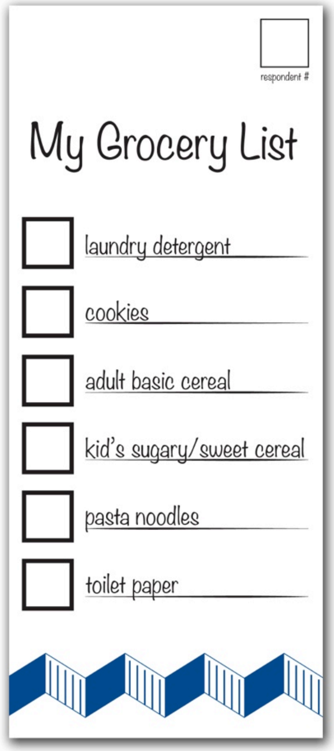


Figure 52: Six Item Shopping List

The participant was asked to enter the realistic and immersive CUshop™ shopping environment. Verbal instructions were provided, where each person was asked to shop as they normally would shop. Participants were asked to shop for six items, seen above through the six item shopping list (Figure 52). The list reads as follows: laundry detergent, cookies, healthy cereal with whole grains, sugary cereal, pasta, and toilet paper. Different than the pilot study, this version asks for participants to pick two cereal types. By adding another cereal item to the shopping list, the consumer is encouraged to spend more time on the cereal aisle where hopefully increasingly reliable data may be collected. After the shopping task, the participant was able to remove the eye tracking glasses and EDA sensor. A post-experiment survey followed the shopping task.

Post-Experiment Survey

At the conclusion of the shopping task, each participant removed the Q-Sensor™ and eye tracking glasses and entered a survey room to complete the post-experiment survey. This survey was completed on a computer using the online tool *Survey Monkey*. The survey consisted of 30 questions listed in 4 sections.

The first section assessed the emotional responses of the consumers towards the designed stimuli. Participants viewed an image of the cereal stimuli and answered pictorial questions using the Self-Assessment Manikin (S.A.M.) (M. Bradley & Lang, 1994). This was identical to the mass online survey completed in using the diagrams found in Figure 45.

The second part of the survey included a 5-point Likert scale to rate the consumer's feelings towards the design keywords used to design the package. The following keywords were used to design the positive package: *sustainable, all-natural, sunshine, happy, healthy, recycled, smiling, green, yellow,* and *blue*. The positive package used graphic design elements that appeared to be happy and healthy while the negative package implemented contrasting graphics. Keywords including: *hazardous, warning, tetrachloroethane, harmful, dangerous,* and *black,* were used to influence the negative package's design. The 5-point Likert scale (extremely negative, negative, neutral, positive, extremely positive) was used in the survey to rate the design keywords.

This self-report measure was taken to validate and describe the valence of any emotional responses recorded through EDA.

The third section of the survey was implemented to understand if the participant was an introvert or extrovert. 13 simple questions were used to understand the character of the person (Cain, 2013). Introverts tend to be less social and quiet, while extroverts are the opposite. Extroverts tend to be more outgoing and social, striving in larger groups. This personality trait was investigated to see if there was a difference in the amount of EDA peaks between introverts and extroverts.

The fourth and last section of the survey asked for basic demographic information and two questions to understand the quality of the participant's vision. The full survey can be viewed in Appendix H.

Post-Experiment Interview

Immediately following the shopping task and survey, a post-experiment interview was conducted to gather additional self-report data (LaBarbera & Tucciarone, 1995). This self-report data was useful in determining the valence of the arousal experienced inside the shopping context. While the participant was completing the post-experiment survey, the experiment proctor quickly uploaded the eye tracking video and EDA data stream. To synchronize the data, a mark was placed in the Tobii Studio software at the exact instance at which the participant pressed the Q-Sensor™ button, as viewed through the eye tracking video. The marker on the eye tracking video was aligned with the marker placed in the EDA data stream viewed within the Q-Sensor™ software. Once the data was synchronized, significant peaks in the data were marked. Once the participant entered the interview room, he/she was shown a play-back video recording of their eye movements during moments of high arousal. The participant was asked to share how he/she felt and any thoughts which he/she remembered at the moment in the video (Guan et al., 2006). This interview was the last step in the EDA consumer study. The participant was thanked for their time and provided proof of participation to receive the incentive of extra credit.

EDA Data Normalization

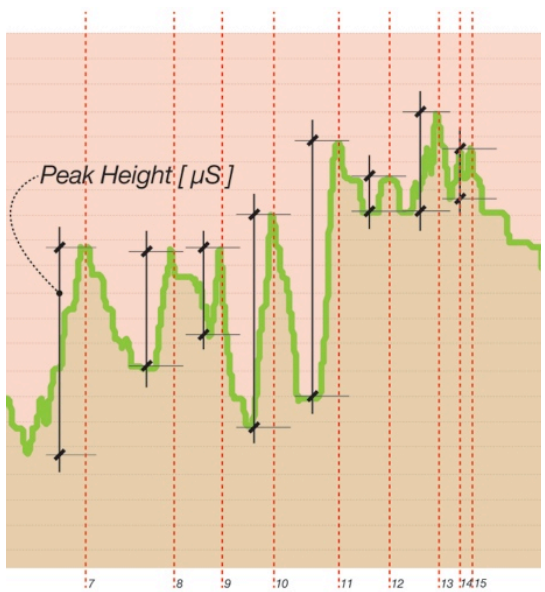


Figure 53: EDA peak height (μS, microSiemens)

An EDA arousal ratio is calculated per stimuli (Figure 53). The following equation is used based on a person's range of extraneous individual differences. The individual's EDA can be expressed as a proportion of their individualized range. This calculation normalizes the EDA data from person to person (Dawson et al., 2007; Lykken & Rose, 1966):

$$EDA\ arousal\ ratio = \frac{(EDA_{peak} - EDA_{min})}{(EDA_{max} - EDA_{min})}$$

$$EDA_{peak} = EDA\ recording\ at\ the\ peak's\ maximum$$

$$EDA_{min} = overall\ personal\ minimum\ EDA\ recording$$

$$EDA_{max} = overall\ personal\ maximum\ EDA\ recording$$

This is a measure of arousal calculated by viewing the peaks within a +5 seconds interval from the first and last fixations of the stimuli (Figure 53) (Latulipe et al., 2011). This number does not describe the valence of the arousal, whether positive or negative. The emotional reactions were confirmed with post-experiment interviews and S.A.M. self-report measures.

Statistical Analysis

EDA arousal ratios

The electrodermal activity of the participants was used to illustrate the arousal levels of the participant throughout the study. Peaks (Skin Conductance Responses, SCR) were located using software provided by the manufacturer of the *Q-Sensor*^{TM} called, *Affectiva Q Analytics.* The software analyzed the EDA outputs (μS, microSiemens) of each participant, using peak detection to report the time, height (EDA), and duration of any abrupt increases in the skin's conductance (Figure 53). An EDA arousal ratio is calculated per stimuli through the equation in the previous section. The individual's EDA can be expressed as a proportion of their individualized range. This calculation normalizes the EDA data from person to person (Dawson et al., 2007; Latulipe et al., 2011). EDA data from the *Affectiva Q Analytics* is entered into Microsoft Excel where a mean EDA ratio is calculated for each participant. StatsPlus, from Microsoft Excel, is used to run a one-way analysis of variance test to see if there is a significant difference in arousal ratios between the participants who viewed the positive or negative stimuli. A p-value less than or equal to 0.05 was considered to be significant (95% confidence interval).

Eye Tracking Metrics

Figure 54: Area of Interest and Analysis

The three eye movement metrics investigated here are *Total Fixation Duration, Time To First Fixation, and Fixation Count*. *Total Fixation Duration* can be described as the total amount of time (seconds) during which a participant spent looking at an AOI, or Area of Interest. How quickly the package caught a participant's eye can be described with *Time To First Fixation*. TTFF is the amount of time, in seconds, for the viewer to make the initial fixation on the AOI, the packaging design stimulus. *Fixation Count* is the number of fixations where the participant fixated on the packaging stimulus, inside the AOI. These three eye movement metrics were exported using Tobii Studio, the software compatible with the mobile eye tracking glasses. StatsPlus, from Microsoft

Excel, was used as the statistical program to run one-way analysis of variance tests in the three categories of TTFF, TFD, and FC comparing the columns of the data from the positive and negative package stimuli. A p-value less than or equal to 0.05 was considered to be significant (95% confidence interval).

Data Correlations

StatsPlus was used to determine if there were any correlations between the EDA data and the S.A.M. scales, and also between the EDA data and the eye movement metrics (TTFF, TFD, and FC). A p-value less than or equal to 0.05 was considered to be significant (95% confidence interval).

Introverts and Extroverts

Self-report data from the post experiment surveys provided information on the character of the participant to determine introverts from extroverts. These questions were asked to see if there was a significant difference between the EDA arousal ratios of introverts and extroverts. The findings from this measure help answer our question on whether or not extroverts are only suitable for EDA consumer testing. StatsPlus was also used to run a one-way analysis of variance test to compare the two columns of data (column A = introvert EDA ratios, column B = extrovert EDA ratios). A p-value less than or equal to 0.05 was considered to be significant (95% confidence interval).

CHAPTER SIX

RESULTS AND DISCUSSION

A total of 42 respondents participated in the thesis study involving electrodermal activity, eye tracking, and the Self-Assessment Manikin. Five of the participants were removed from the final data set due to poor calibration to the eye tracking glasses. All of the participant eye tracking videos were viewed to confirm that each person fixated on the emotional stimulus presented. One of the participants was removed because they did not fixate on the test variable. After a total of six participants were removed, the residual left over was calculated to be 36 participants, split evenly between the positive and negative test variables. Additional demographic information is summarized and presented in Appendix B.

One-way analysis of variance tests were completed for the following results to report significant findings. These ANOVA tests compared the means of two samples using the F distribution. A 95% confidence interval was used to record a finding as significant. Those that produced p-values less than 0.05 were considered to be significant. StatsPlus for Microsoft Excel was used to execute the one-way ANOVA tests recorded in Appendix F.

Self-Report Interview Responses

Inside the Affectiva Q software, comment markers were placed to record the verbal responses from the interview section of the study procedure. These verbal responses were transcribed by the experiment proctor and saved to the EDA file. Below in the following table, several responses are shared to highlight the trends found in the interview portion of the study. The participant was asked to respond with personal thoughts in relation to each video clip. This interview data was used to confirm any emotional reactions to the packaging stimuli or the experimental design.

TABLE 01			
Examples of Transcribed Interview Responses			
Viewed the Negative Stimulus		Viewed the Positive Stimulus	
"I noticed the size differences between the different packages across the cereal aisle. The generic was placed on the bottom. I obviously noticed the cockroaches on the package [negative stimulus design]. This made it feel out of place and very unexpected."	*"Here I was just trying to figure out which cereal I liked the best, which one really suited me…The one with the girl was very abnormal, it made me feel weird."*	*"Emotionally, I was not doing anything. I was looking at quantity, looking for the best value. With the kid's cereal, I just picked something that I liked. I picked the Kroger because I'm from Atlanta, it is normal to me. Healthy Oats was funny, it looks cheesy and out of place."*	*"Here I was picking an adult cereal. I picked something that I knew had health benefits. When picking children's cereal, I looked for something that I knew from Harcombe, like Reese's Puffs. I was not really surprised with the Healthy Oats, because claims like healthy and heart wholesome don't always resonate until you read and check the back of the box."*

Table 1: Examples of Transcribed Interview Responses

At first look, the interview responses appear to highly suggest that the viewers did not view, read, and comprehend the emotional stimuli presented. The interview question was, in some degrees, left open to be vague with the intention of not forcing a certain answer from the participant. The respondents were shown clips of their eye tracking movements through the cereal aisle and asked to respond with any emotional reactions they experienced during the video. Most respondents simply provided answers that explained the reasoning behind picking their favorite cereal in the aisle. Few of the respondents reported viewing the emotional stimuli. This could be attributed to two reasons. First, before the shopping task, the respondent was asked to "shop as they would normally shop." The quick times spent on the cereal aisle and the low verbal reporting of viewing the emotional stimuli suggest that participants were visually searching for their favorite brands within the cereal aisle. Little to no consideration was given towards packaging designs that were unfamiliar to the consumers. Therefore, the emotional stimulus, new to the consumer, could have been simply disregarded. Secondly, the low reporting suggests that the interview question provided was too vague of a question. Trends show that the consumers were just providing verbal reasoning for why they picked certain brands. Possibly, the consumers felt that they did not need to verbally explain reasoning for not picking an item. When asked directly if the consumer viewed the stimuli within the store, the participant would admit to seeing the packaging. This answer could have been forced though out of fear of answering an interview question "wrong" in the perspective of the participant. The discrepancies found with these self-reporting measures supports the foundation of an investigation in physiological research. It is the job of the physiological recording device to reveal how people really feel and not the cognitive efforts of the participant. Often people don't necessarily realize what they feel.

Data provided from the Q-Sensor™ and the eye tracking glasses was still analyzed for the 36 participants that calibrated correctly and visually fixated on the emotional stimulus within the shopping context. Reasoning for continuing the analysis of participant's data that did not verbally report seeing the emotional stimuli can be defined by searching for trends that could appear to be subconscious, emotional directives which can only be captured by physiological recording devices, instead of relying on the misdirected conscious strings of self-report interview data.

Electrodermal Activity

Once stimulus time intervals were established for each participant, the peak detection files were read to understand if any peaks were experienced during the viewing of the stimulus. Table 2 summarizes the peaks experienced per participant. An obvious trend was found, reporting that almost all of the participants experienced a total of zero peaks per stimulus interval. Only three out of 36 participants (8.33%) that viewed the emotional stimuli experienced peaks in electrodermal activity. It was determined that these three participants be understood as outliers in the data because they each experienced a very high level of peaks (91, 48, and 17), while the larger trend were found to have zero peaks. When compared to the peaks experienced in the pilot study and showing an increase in sample size, this supports the reasoning that EDA is very difficult to apply to packaging studies within the shopping context.

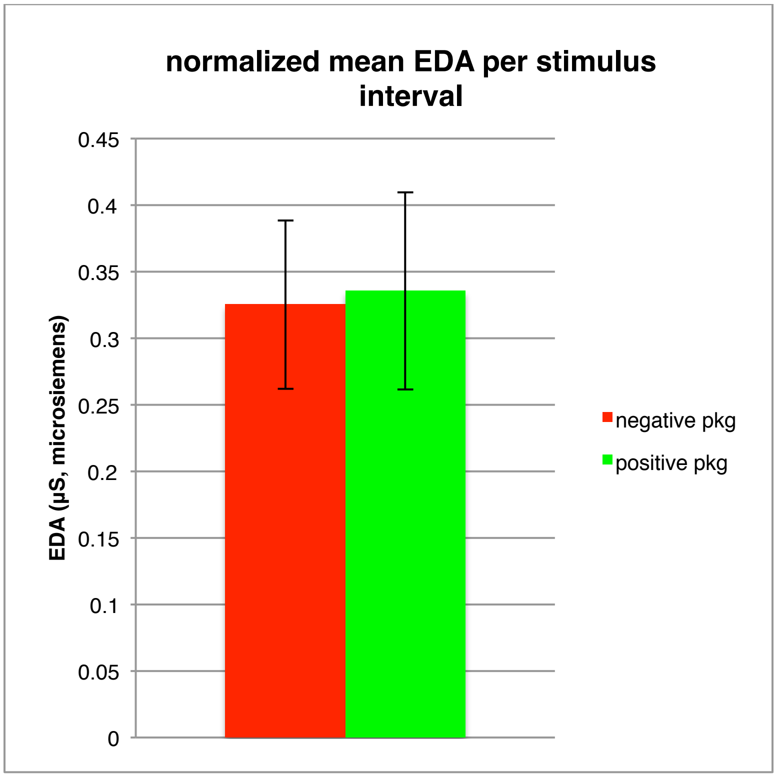


Figure 55: Mean EDA Per Interval

The intent of the methodology states that a mean EDA would be established for each of the emotional stimuli based off of EDA recorded at the peaks of each participant. This could not be generated due to the fact that the majority of participants did not experience any peaks within the interval. Figure 55 was generated as an alternative graph to investigate if there was any trend in EDA per stimulus interval. No significance was found between the two stimulus groups.

The participants of this thesis study did not experience peaks, while conversely peaks were experienced in the pilot study. The one clear difference between the pilot study and the thesis study is the implementation of the large quantity of white, neutral context packages used around the pilot study stimuli. Peaks reported in the pilot study could be due to physiological reactions fueled by the strange appearance of the large quantity of white, neutral cereal packages. Before the thesis study, it was concluded that the pilot study peaks were connected to the emotional stimuli. However, the low arousal ratios from the participants of this thesis study support the reasoning that the pilot study peaks were caused by the strange, white packages. This intersection between the pilot study and the final study highlights a fundamental issue in the experimental set-up of stimuli in shopping contexts. It is very difficult to naturally engage the consumer in experiencing the emotional stimuli within a shopping context. While a shopping context will produce natural, subconscious reactions to packaging, the quick and non-linear movements of consumers make recording and understanding these reactions very difficult.

Introverts and Extroverts

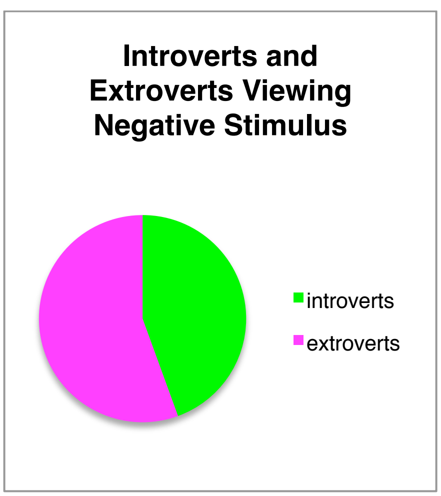


Figure 56: Personality that Viewed Negative Stimulus

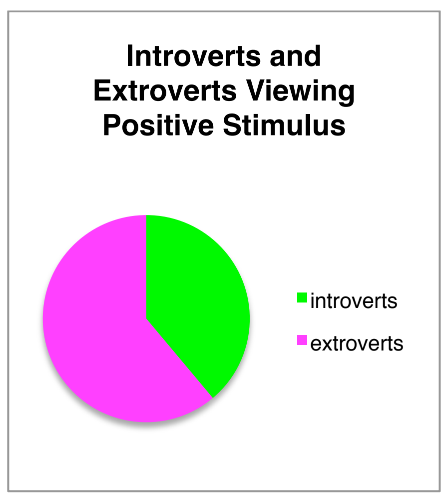


Figure 57: Personality that Viewed Positive Stimulus

Part of the post shopping task survey asked participants 11 true or false questions to label the consumer as an introvert or extrovert. Figure 56 - Figure 58 illustrate that there was a small majority of extroverts that participated in the study. Due to the fact that the EDA data provided insignificant results, we could not investigate if there was a trend between the EDA data of introverts and extroverts. The large majority of participants did not experience EDA peaks during the stimulus intervals. Personal character did not influence the EDA arousal ratios.

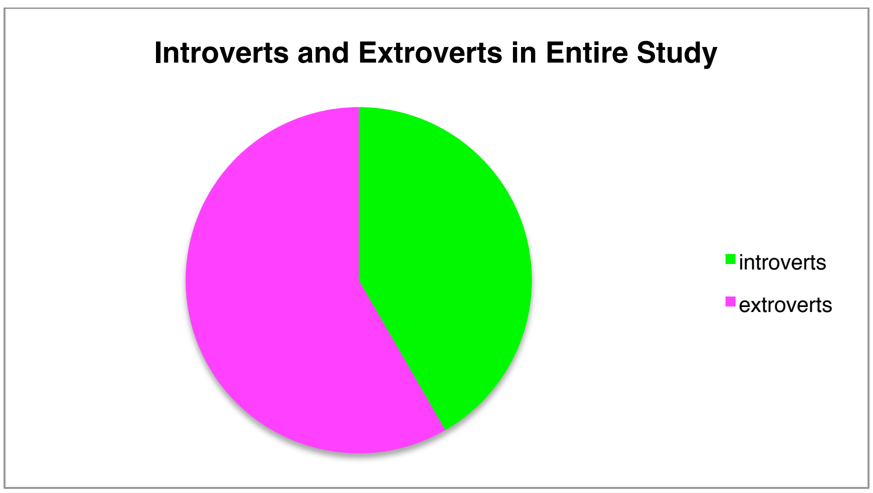


Figure 58: Personality through Entire Study

Eye Tracking

The two images below are aggregate heat maps exported from Tobii Studio (Figure 59 - Figure 60). The images highlight the "hot-spots" across the Area of Analysis (AOA) in areas of red. These are useful in understanding any visual patterns experienced by the entire group of participants. No major visual trends are shown across the packaging stimuli. However, the top right of each image is somewhat weighted. It can be hypothesized that these large areas of visual attention could be linked to the solid and brightly colored, orange packages. An example of two participant's scan paths across the Area of Interest (AOI) is provided in Figure 61.

Figure 59: Aggregate Heat Map, Negative Stimulus

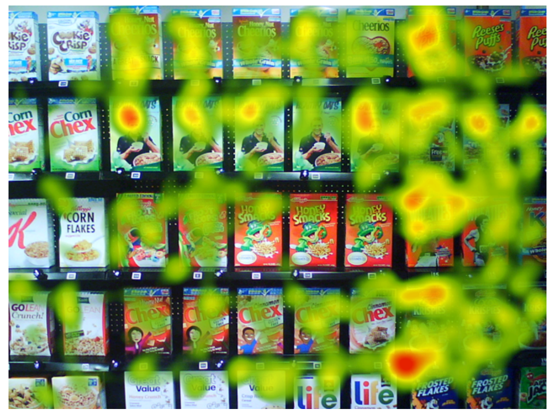


Figure 60: Aggregate Heat Map, Positive Stimulus

Figure 61: Example of Two Participant's Scan Paths

Time to First Fixation (Figure 62), Total Fixation Duration (Figure 63), and Fixation Count (Figure 64) was investigated to see if there were any significant differences in eye movements per the two stimuli. These metrics apply to the Area of Interest (AOI), labeled as the shelf area containing the emotional stimuli. The detailed reporting of each respondent's eye tracking metrics can be read in Appendix E.

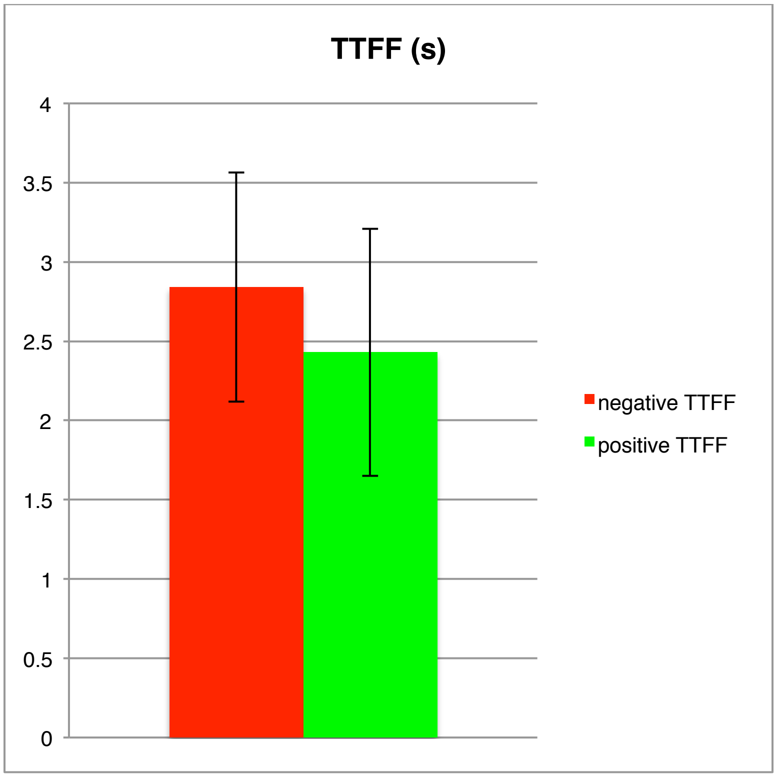


Figure 62: Time to First Fixation (TTFF)

Time To First Fixation (TTFF) reports the amount of time, in seconds, which it took for a participant to fixate on the Area of Interest (AOI). There was no significant difference ($p>0.05$) between the Time To First Fixation of the positive and negative stimuli (Figure 62). This supports one of the conclusions of the pilot study. The pilot study reported increased TTFF for packages with minor graphics and little package design. Those packages that were designed with color and graphics reported smaller TTFF. Due to the fact that all of the packages in the thesis study were designed with graphics, it is expected that there is no significance between the two packaging stimuli.

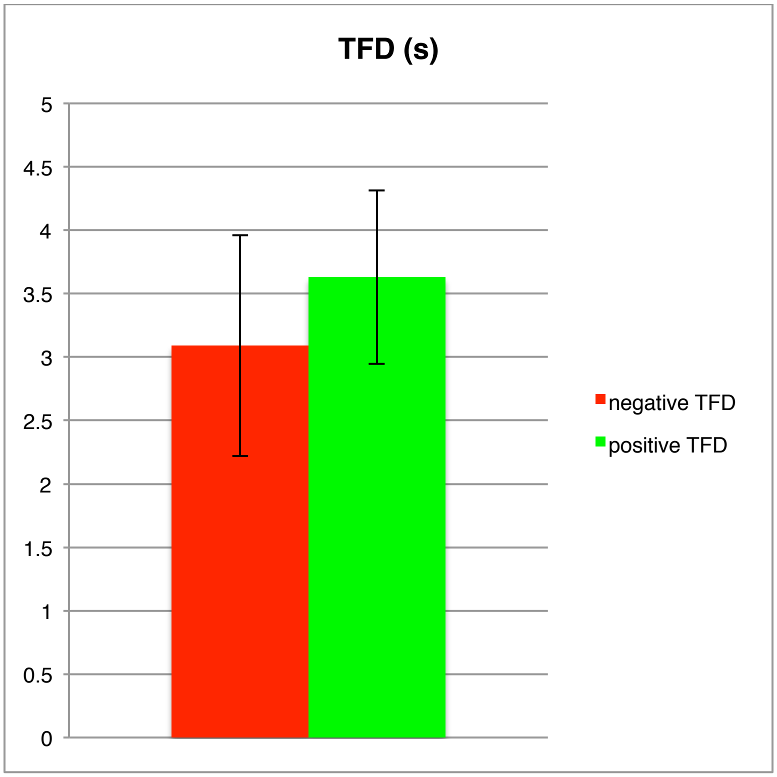


Figure 63: Total Fixation Duration (TFD)

Total Fixation Duration (TFD) defines the total amount of time, in seconds, during which the viewer provided fixations along a particular Area of Interest (AOI). There was no significant difference ($p>0.05$) between the Total Fixation Duration between the positive and negative stimuli (Figure 63). The lack of significance between the two groups' fixation durations highlights the reasoning that both groups spent an equal amount of time viewing the stimuli before making a purchase decision. Both groups viewed the two different stimuli equally.

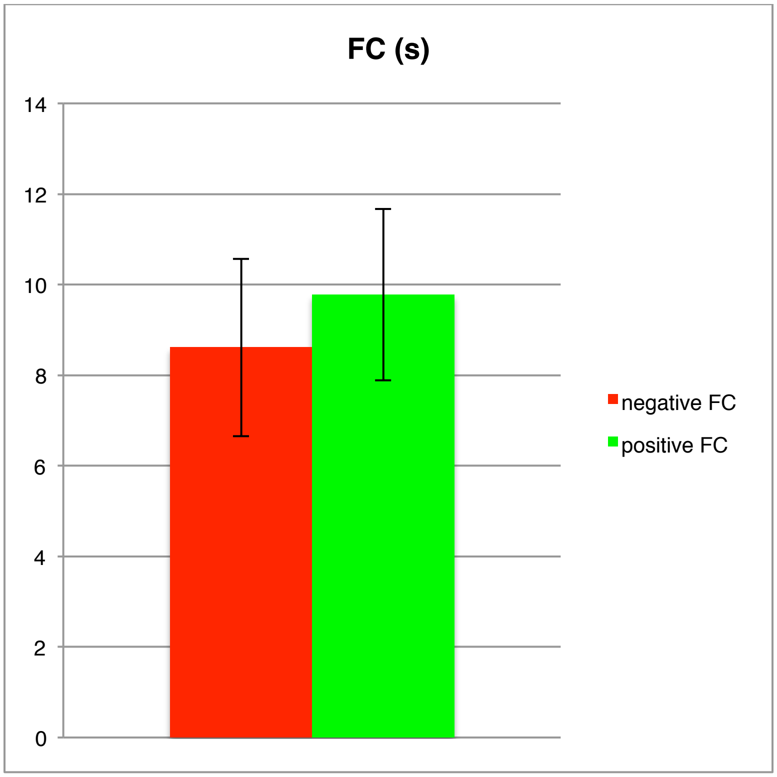


Figure 64: Fixation Count (FC)

Fixation Count (FC) is an eye tracking metric that reports the total amount of fixations on the Area of Interest (AOI). There was no significant difference (p>0.05) between the Fixation Counts of the positive and negative stimuli (Figure 64).

The Self-Assessment Manikin

The Self-Assessment Manikin (S.A.M.) was employed in two parts. First, the self-report measure was the main part of the post shopping task survey completed by all thesis study participants. If peaks had been recorded across the participants during the stimulus interval, this information would have been applied to the EDA data to determine the valence of the arousal, whether it was a positive or negative emotion. The S.A.M.

survey provides the link between arousal and descriptive emotion. Secondly, the S.A.M. was used as a checking device to ensure that the designed stimuli were actually positive or negative. This was completed with a large online survey consisting of 126 respondents, collected and incentivized by the online tool *Survey Monkey*.

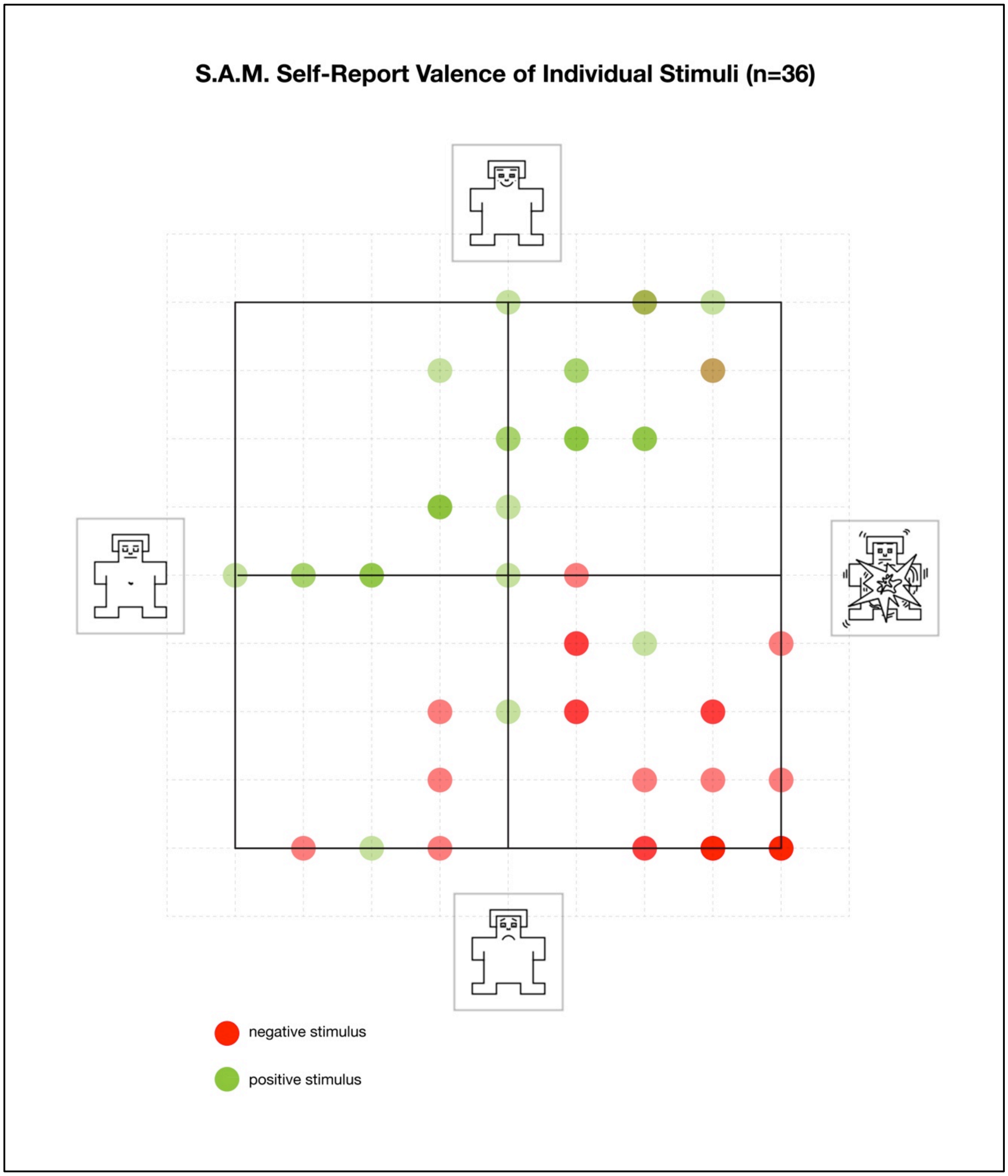


Figure 65: S.A.M. Self Reported Valence, small sample

Figure 65 illustrates the two main emotional factors of the S.A.M. test in an x,y coordinate system. The x-axis denotes boring or low-stimulated responses in the negative range and exciting or highly stimulated responses in the positive range. The y-axis denotes sad or depressing emotions in the negative range and happy or pleasurable emotions in the positive range. The red dots illustrate that participants found the negative package design to be highly arousing and highly depressing or sad. The green dots show an opposite trend of highly pleasing or happy feelings with a more neutral arousal representation.

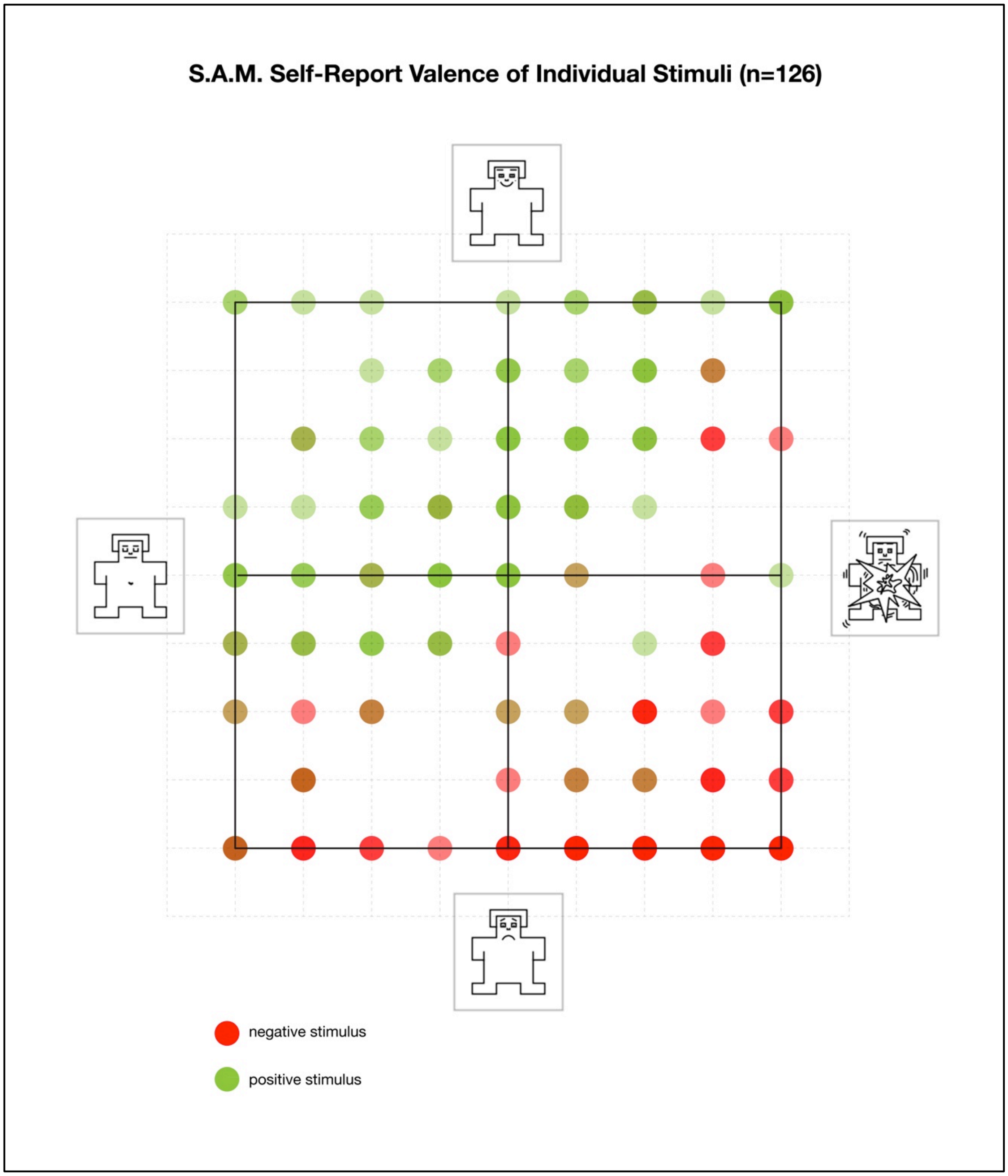


Figure 66: S.A.M. Self Reported Valence, large sample

Similar responses were collected from the online survey, illustrated in Figure 66. The figures that follow show the significant differences found between the two package designs. Both samples that were surveyed reported similar responses, showing significant results. This significance supports the correct presentation of the stimuli in engaging the consumers with negative and positive emotions.

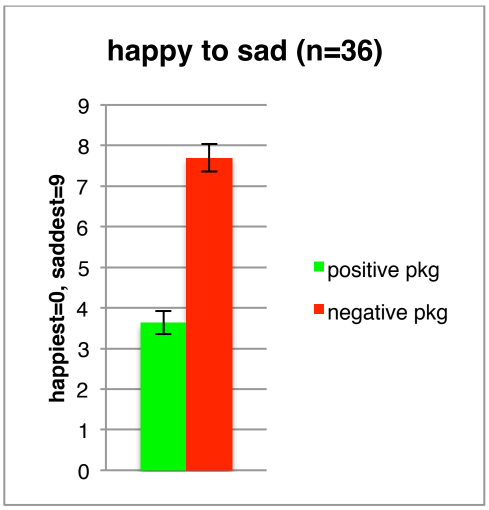


Figure 67: S.A.M. Happy to Sad Scale, small

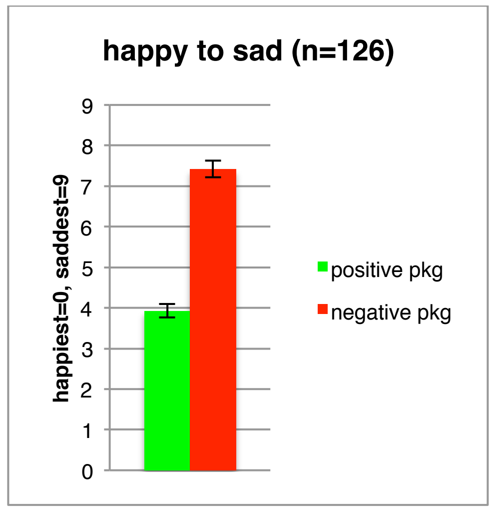


Figure 68: S.A.M. Happy to Sad Scale, large

The happy to sad range of the y-axis of the S.A.M. produced significant results when comparing the negative and positive designed stimuli (Figure 67 - Figure 68). Both surveys (n=36 and n=126) produced p values less than 0.01. The positive package was rated as being a happier and more pleasant design compared to the negative package that was rated as being sad and more depressing.

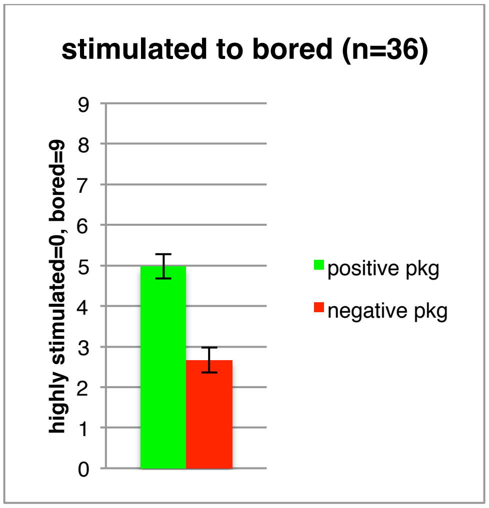


Figure 69: S.A.M. Stimulated to Bored, small

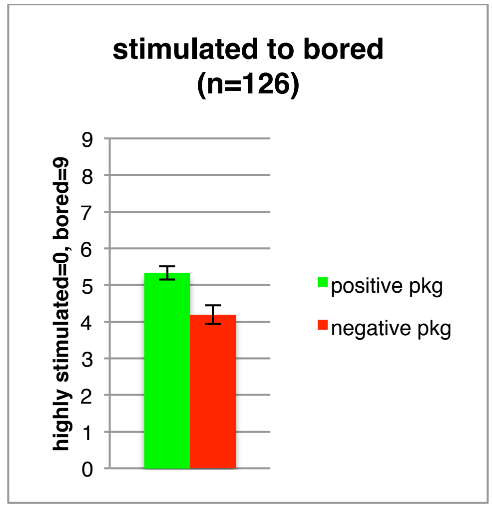


Figure 70: S.A.M. Stimulated to Bored, large

The stimulated to bored range of the x-axis of the S.A.M. produced significant results, both resulting in *p* values less than 0.01 (Figure 69 - Figure 70). The positive package design was rated as a neutral design, situated half-way between highly stimulated and bored. The negative design was rated as a more arousing or stimulating design when compared to the positive design. This supports the experimental methodology in that the negative design could be a more arousing stimuli that a happy, positive design. Negative emotions are usually more arousing than positive emotions.

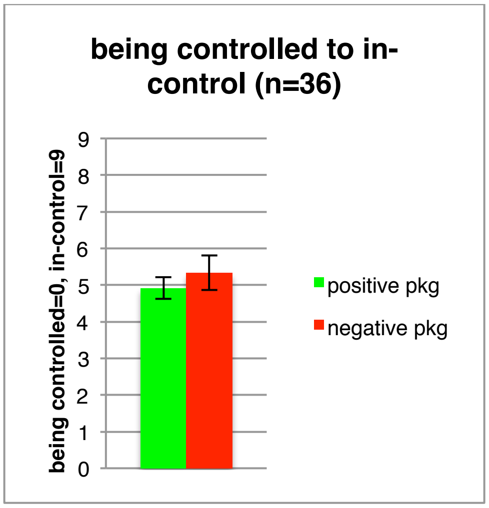


Figure 71: Controlling, small

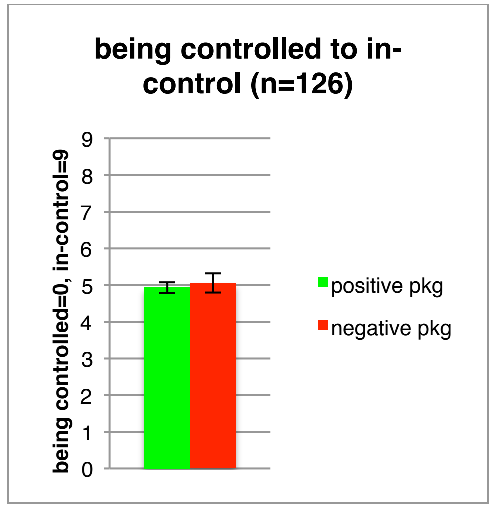


Figure 72: Controlling, large

The third and last part of the S.A.M. produced insignificant results, both resulting in p values greater than 0.05 (Figure 71 - Figure 72). This scale ranged on feelings of being controlled to being in-control. It is hypothesized that the insignificance of this scale between the two package designs could have been caused by the strange request of the question of dominance. The values that were produced are listed as mid-range, which describe the neutral response to the question.

Design Keywords

The design keywords listed in the methodology were tested on a 5-point Likert scale to determine the emotional valence of the keyword. The range surveyed read as follows: *extremely negative, negative, neutral, positive, and extremely positive*. The design keywords were divided into two separate groups, defining the attributes of the negative and positive package designs. This survey was applied to the thesis study survey of 36 respondents as well as the online survey consisting of 126 respondents. A significant difference was found in both surveys (p<0.01) when comparing the positive and negative design keywords (Figure 73 - Figure 74). The positive design keywords were rated closer towards the positive end of the Likert scale, while the negative design keywords were placed towards the opposite end. This survey question was implemented as an additional checking device in supporting the valence of the packaging design stimuli.

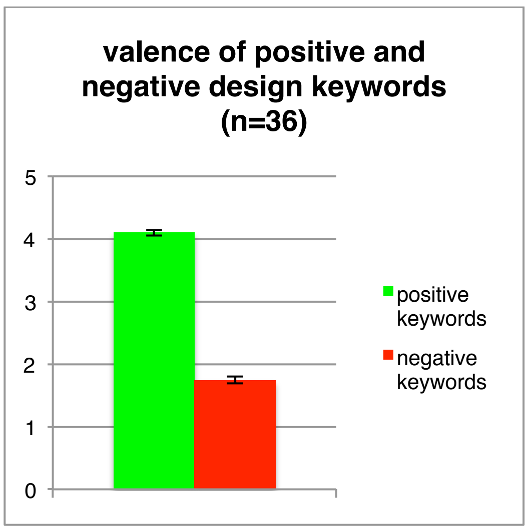


Figure 73: Valence of Positive and Negative Keywords, small sample

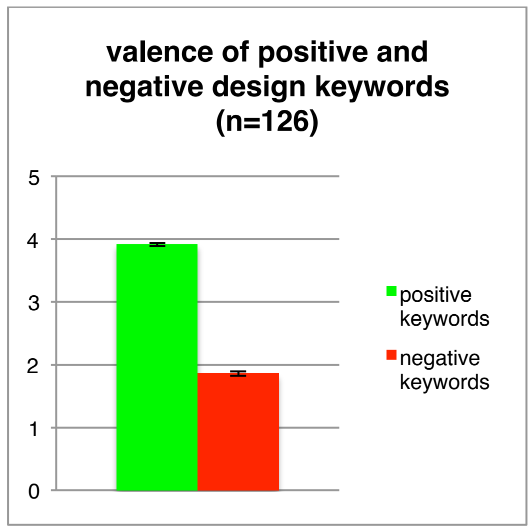


Figure 74: Valence of Positive and Negative Keywords, large sample

The following figures (Figure 75 - Figure 76) illustrate the responses to the design keywords ungrouped. A large drop can be visualized between the keywords used to design the positive and negative packages.

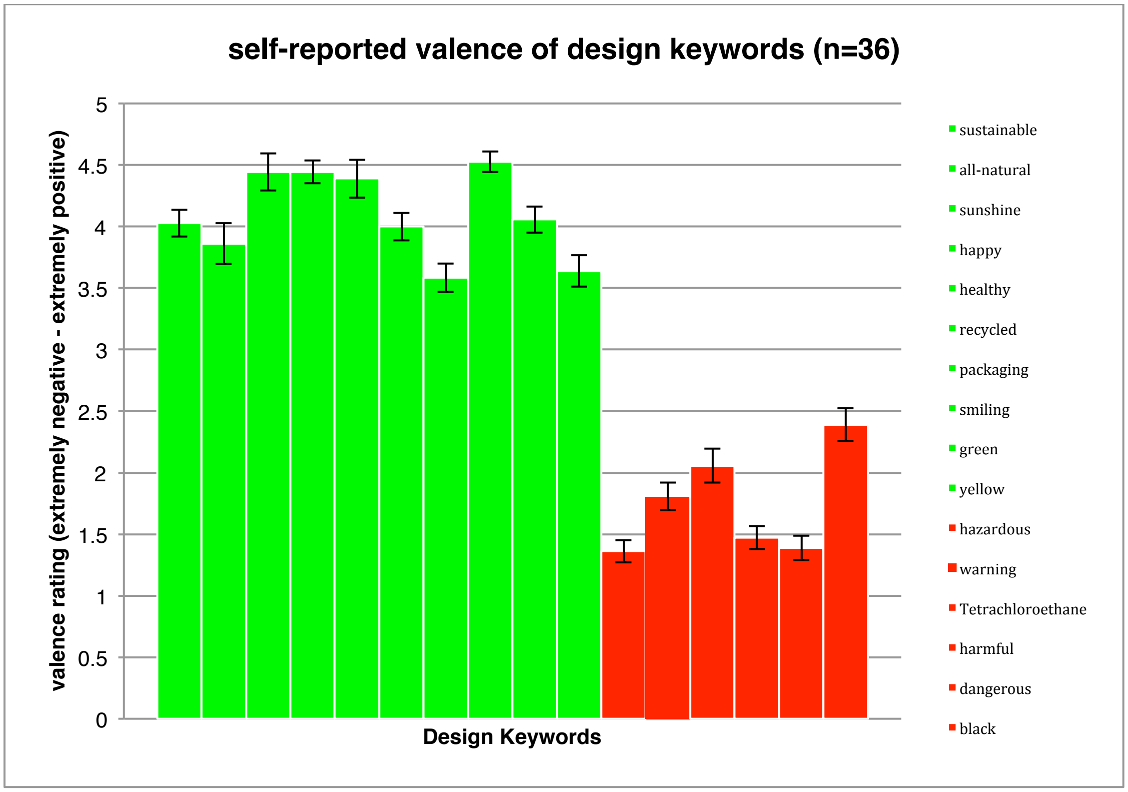


Figure 75: Valence of Individual Design Keywords, small sample

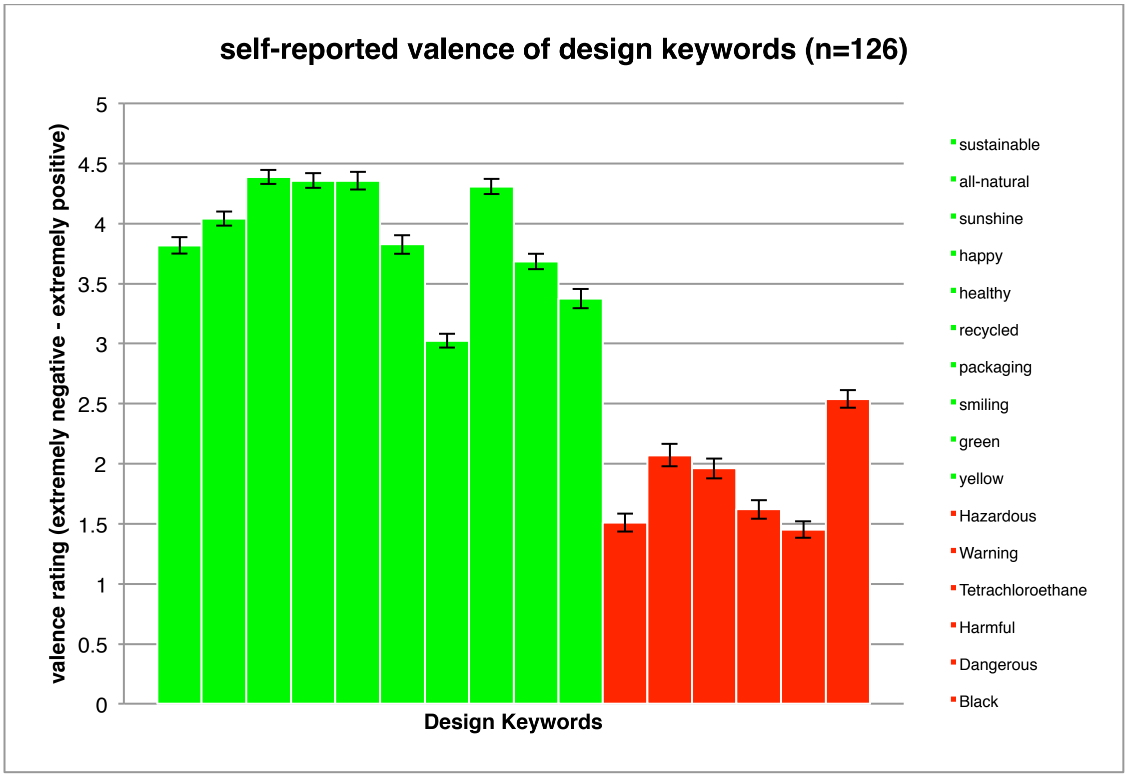


Figure 76: Valence of Individual Keywords, large sample

Data collected from physiological and self-report tools of the study support the decision to fail reject the null hypothesis described in the experimental methodology. Under the circumstances of the pilot study and full-scale study we did not find significance, however we cannot conclude that EDA is not a valid tool for measuring arousal towards packaging designs inside the shopping environment. Emphasis in the previous statement should be placed on the shopping context. Many changes were made to the methodology of this thesis study to respond to the recommendations of the pilot study. Improvements were made such as: analyzing the valence of the designed stimuli with non-verbal pictorial images, increasing the sample size to at least 30 participants, activating a higher maximum EDA level from the consumer with a singing activity in the calibration test, removing the neutral designed package for intentions of clarity, and increasing the time spent viewing the cereal aisle with the addition of another cereal item to the shopping list. The technological advances of EDA in consumer studies are admirable, especially in the case of the mobile device of the Q-Sensor^{TM}. Future studies should closely follow the advancements of EDA in hopes that different experimental constraints could report significance towards arousal and packaging.

CHAPTER SEVEN

CONCLUSIONS

This thesis highlights the quick and fast decisions that consumers make in the shopping context. A subconscious, emotional directive guides consumers throughout stores to complete purchasing decisions. Electrodermal Activity is introduced to the toolset of packaging designers in an effort to better understand the emotions that connect consumers to the products that they buy. This experiment is unparalleled in that it increases the reality of the study participant to the highest extent. The studies presented are significantly different from the few past studies completed with EDA and packaging design (Groeppel-Klein, 2005; Liao et al., 2012). Study participants participate as if they are realistically shopping in the retail environment. This insertion of reality reveals the complexities of consumer behavior where many fast processes occur in short flashes of time.

Under the circumstances of the pilot study and full-scale study we did not find significance, however we cannot conclude that EDA is not a valid tool for measuring arousal towards packaging designs inside the shopping environment. However, this is not to be stated in complete abandonment. While a shopping context will produce natural, subconscious reactions to packaging, the quick and non-linear movements of consumers make recording and understanding these reactions very difficult. The review of literature tags the opinions that EDA requires difficult interpretation, however the same technology is also supported with requests for future experimentation and development. A varied list of existing applications involving EDA as a measure of physiological processes illustrates the far reaching efforts of this technology. EDA has seen much success through the efforts of studies involving longer periods of measurement. The conclusion of this study is not meant to disregard all future intersections with EDA and packaging evaluation. Instead, more tactics in experimental design must be used to understand the connection between the brain, the body's nervous system, a consumer's visual attention, and their self-report measures.

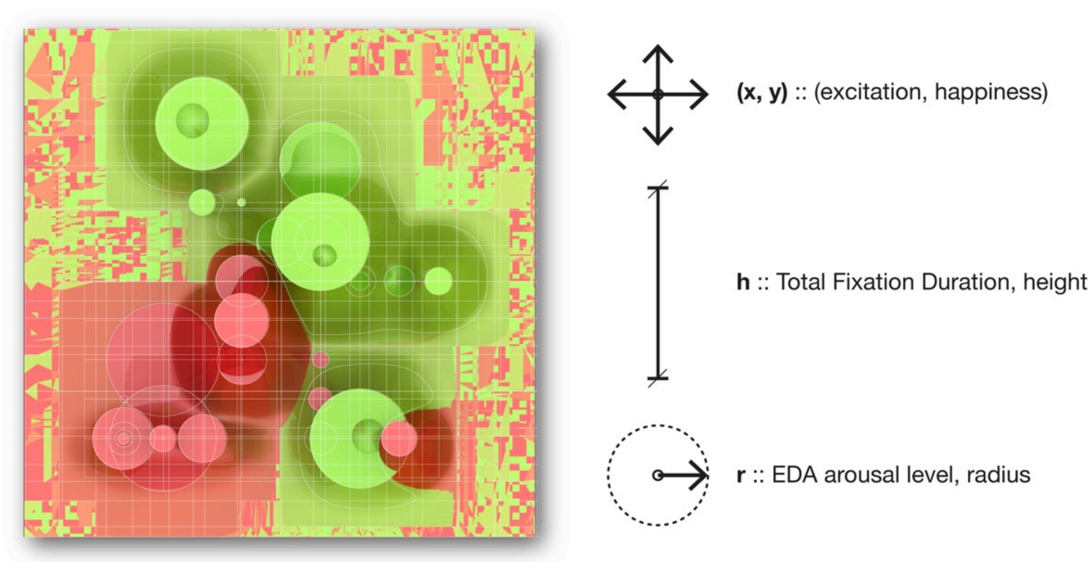


Figure 77: 3-Dimensional Emotion Mapping Concept

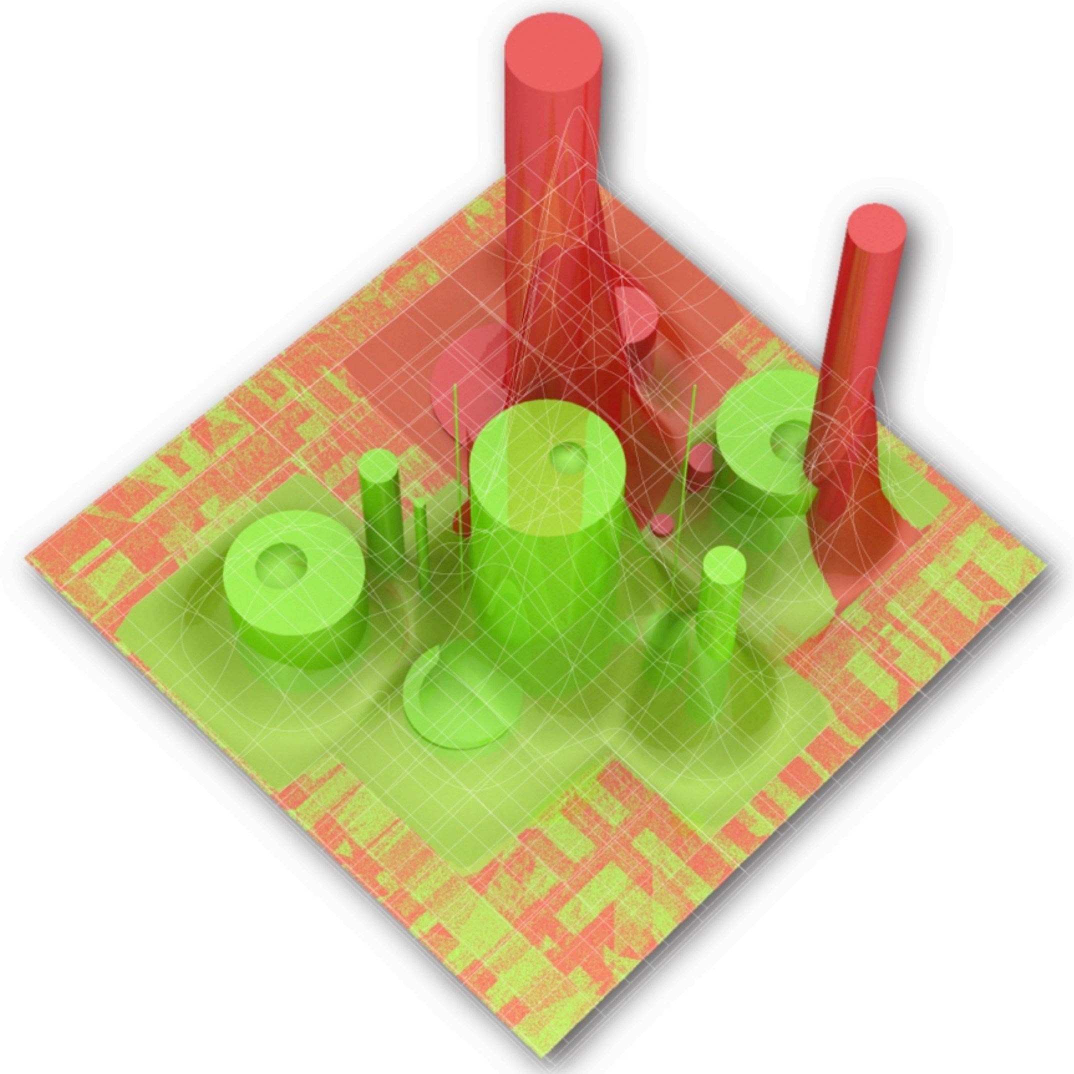


Figure 78: 3-Dimensional Emotion Mapping Conceptual Model

Figure 77 - Figure 78 illustrate 3-Dimensional model renderings of a future of EDA, eye tracking, and S.A.M. visualizations in tandem as a triangulation of datasets. The results taken from this thesis are not ready for an accurate dimensional mapping concept. However, with future advancements in the intersections of these technologies, a physical and tangible experience can be generated to express the physiological responses

of the consumer. To combine the data from the three tools used in this thesis, a multidimensional scaling approach is conceptualized (Rebollar, Lidón, Serrano, Martín, & Fernández, 2012; Wickelmaier, 2003). The S.A.M. scales are plotted as x,y coordinates on the xy plane. The Total Fixation Duration of the participant determines the height of each plotted point. The EDA arousal degree adds the final layer of information as a radius to a circle. Cylinders are lofted between the extruded circles and a surface is generated between the three dimensions of the model. Products can quickly be evaluated and visually compared against one another to understand the degree and valence of a consumer's arousal. The spirit and value of the research of emotions towards packaging design can be experienced through this model form.

Emotional reactions occur during the critical instance of the purchasing decision where emotion is the key to action, not conscious and rational thought. With a better understanding of measuring consumer emotion, designers are provided with a process by which to design better interactions for people. With a powerful understanding of emotion, designers can help target the claims that draw consumer engagement. Theorized dangers such as "the buy button," in the brain have been criticized by neuromarketing journalists in the wake of physiological research in advertising (Blakeslee, 2004; Witchalls, 2004). While a better understanding of a consumer's emotion could result in a negative direction, safe guards should be implemented to encourage designers to use these tools for the benefit of humankind. While we can't expect consumers to make rational decisions in the retail environment, we [designers] can help people make healthy emotional decisions. With the knowledge of a consumer's physiological responses we can be inspired to provide people with positive influences at the point of purchase. This quick time interval, but large influence, will be brought to the home to positively drive a more healthy and happy lifestyle. The merging of this scientific field in packaging design represents the major influence of packaging within our society.

This thesis concludes that additional research must be completed to determine the validity of electrodermal activity as a measure of arousal towards packaging in the shopping environment. A future study could be constructed to use universally liked and hated products, such as chocolate and cod liver oil, to attempt to measure arousal in the shopping context. Price as a factor could be involved in the study to increase participant involvement. Studies should exhaust a list of highly engaging products and consumer environments in order to allow for a definitive conclusion to the validity of EDA towards packaging in the shopping environment.

Emotions are highlighted and valued as an important area of research in consumer behavior. Variables must be carefully isolated so as to not distract from the realism of the shopping environment. Packaging designers should aim to parallel the fields of neuroscience, psychology, and physiology in order to attempt any applications in measuring consumer emotions. By combining the successes of these fields, an exciting and rewarding future of emotion in consumer behavior can result.

CHAPTER EIGHT

RECOMMENDATIONS

This thesis started on a journey to describe the quiet but highly emotional conversations that consumers have with packaging inside the shopping context. Additional attempts in combining these physiological tools should be evaluated before the technologies are completely disregarded and the connection to packaging and emotion is shattered. Several recommendations are provided in this chapter to address different areas of future research. A complete exhaustion of these areas of research will help packaging designers make a highly qualified decision on the weakening connection between packaging design and electrodermal activity within the shopping context.

The study documented in this thesis could be expanded and replicated against different product categories to confirm or challenge further the complete disassociation of electrodermal activity to packaging design evaluation. Universally liked and hated products could be packaged to provide a stronger emotional reaction. In addition to different product categories, the tools could be applied to different environments of ranging types and sizes. In an effort to increase the realism experienced by the study participants, the real price of items within the store could be displayed to help consumers make their purchasing decisions.

Current practitioners that utilize EDA in non-shopping contexts should highly consider the true validity of their results. Typical studies that isolate the packaging stimuli on projection screens and televisions in dark, sterile rooms should be given a close inspection. Placing consumers in unrealistic environments should lead to invalid or unnatural conclusions.

Product branding should also be evaluated and where possible, existing brands should be removed from emotion related studies. If all cereal packages in the study were unfamiliar to the participant, then possibly, the consumer would have responded differently. When placed around unfamiliar products, the consumer could have viewed each of the products equally. Brand recognition within the participants caused new brands like the emotional stimuli to be left unnoticed. To provide more presence to the unknown stimuli, an entirely new store full of products could be generated. Unknown products with no similarities to existing brands could be utilized to generate an evenly competitive market of products.

Additional physiological devices should be added to the packaging evaluation toolset to parallel research in neuroscience and other advancing fields. Tools including, but not limited to, electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and facial recognition should be implemented into the toolset of the packaging designer. Links between these fields of research could reveal more explicit information towards emotion and decision making with packaging. As these fields advance and their tools become more accurate and responsive, these units will hopefully follow trends of becoming wireless and lightweight. These streamlining attributes will make their introduction into the shopping context highly applicable.

When designing emotional stimuli, the experiment designer should establish multiple checkpoints to ensure that the intended valence of the package design is clear. This can include S.A.M. surveys depicting only pictorial scales and Likert verbal scales. An additional tactic could be taken by using images from the *International Affective Picture System*, a set of highly validated images. These images could be used to define the emotion of the package before testing as a control of both rated pleasure and arousal from viewers (M. M. Bradley et al., 2008). The far reaching effects of emotion in consumer behavior presents an appropriate need for further research in the recording of physiological responses to stimuli. An exciting future is hopeful for packaging designers that share a common interest with the fields that merge the common grounds of neurology, psychology, and physiology.

APPENDICES

Appendix A

Demographic Summary: Pilot Study

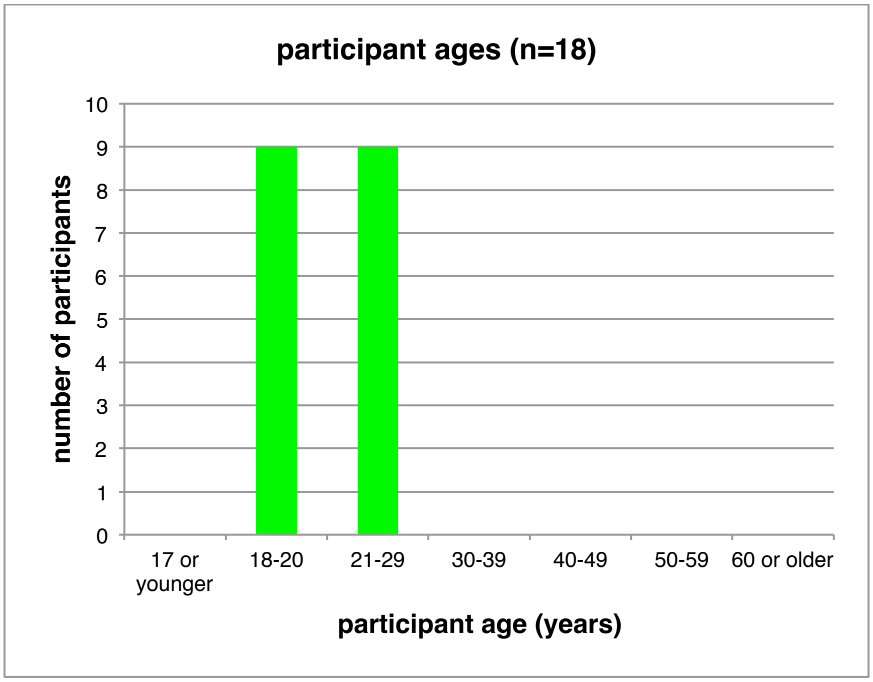


Figure 79: Participant Ages, Pilot Study

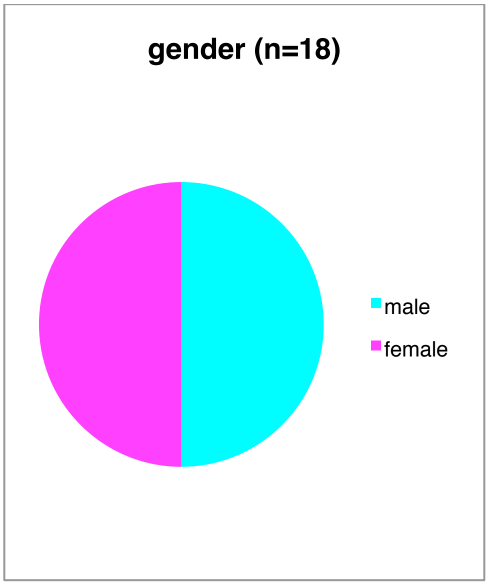


Figure 80: Participant Gender, Pilot Study

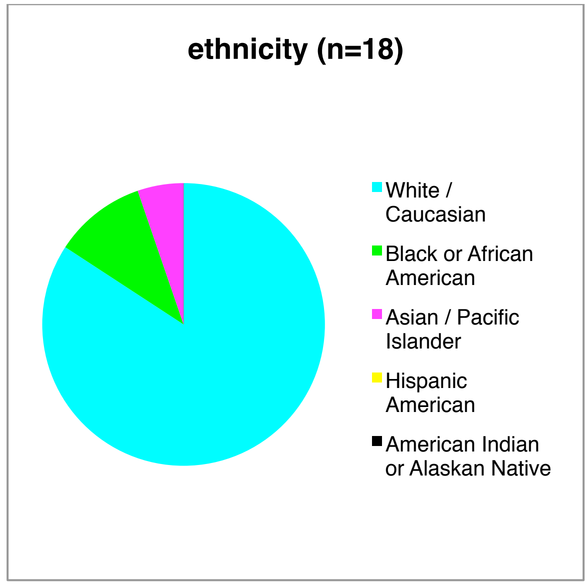


Figure 81: Participant Ethnicity, Pilot Study

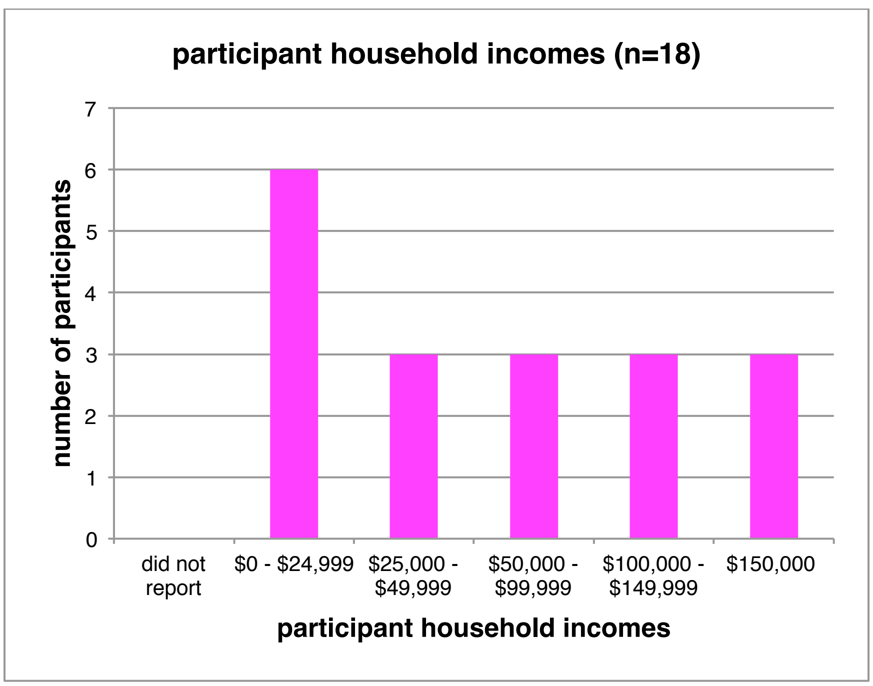


Figure 82: Participant Household Incomes, Pilot Study

Appendix B

Demographic Summary: Thesis Study

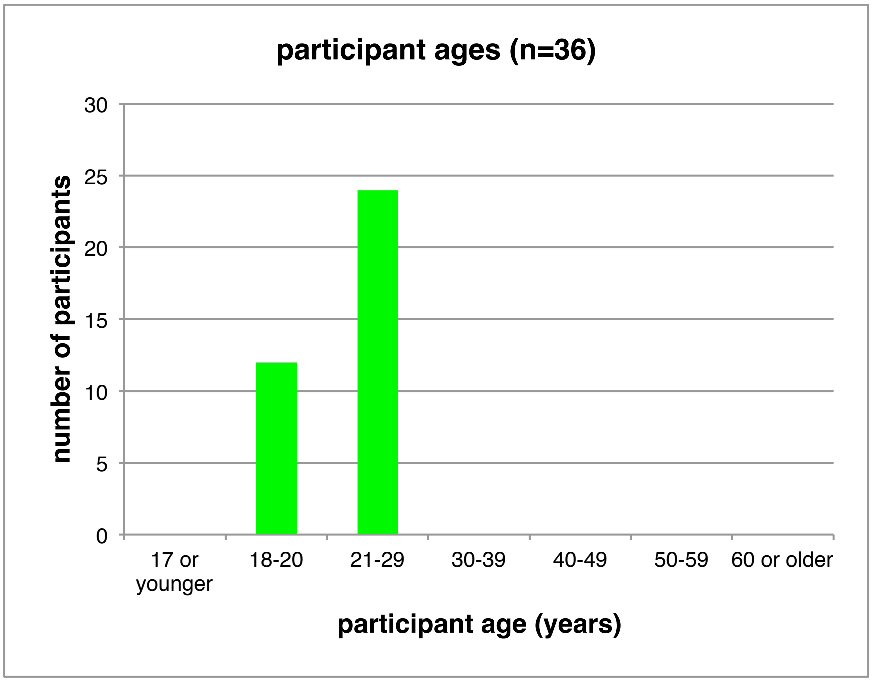


Figure 83: Participant Ages, Thesis Study

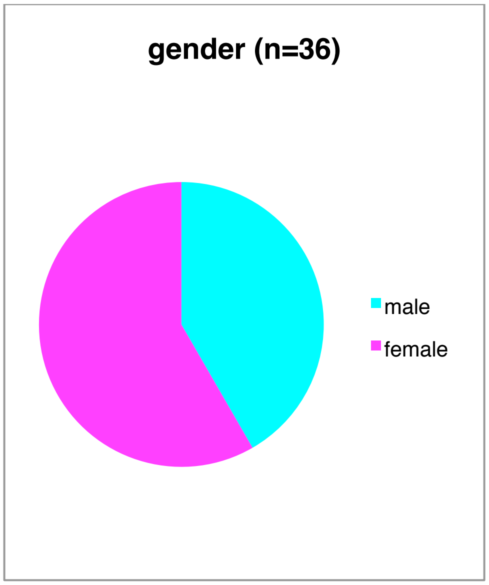


Figure 84: Participant Gender, Thesis Study

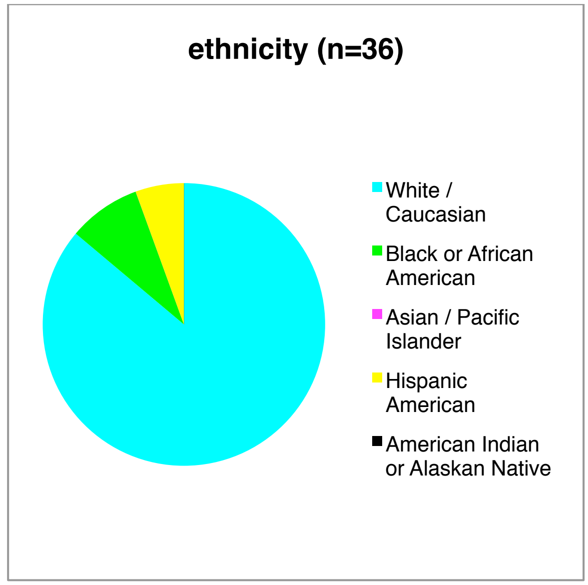


Figure 85: Participant Ethnicity, Thesis Study

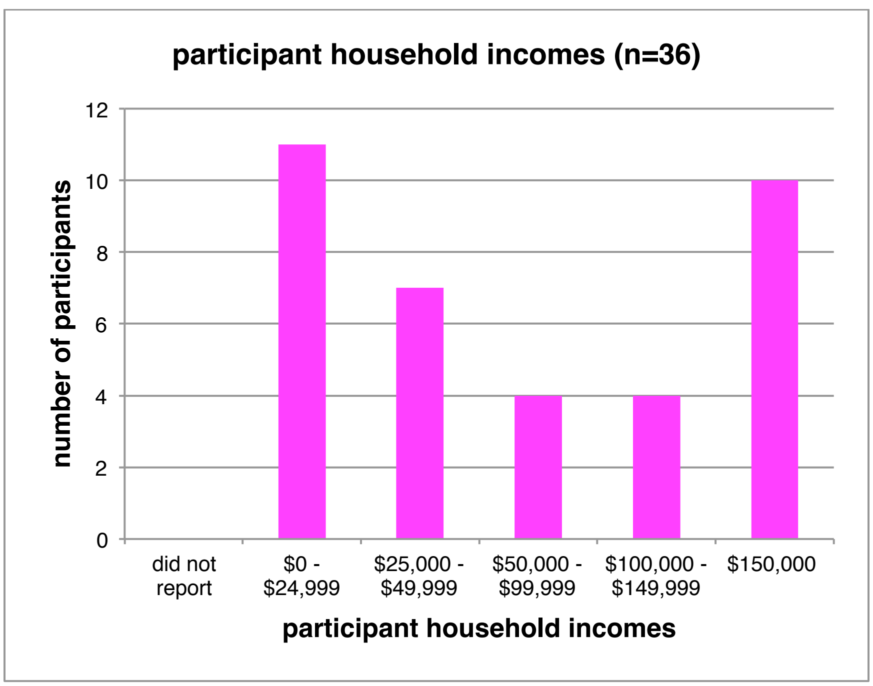


Figure 86: Participant Household Incomes, Thesis Study

Appendix C

Demographic Summary: Online Survey

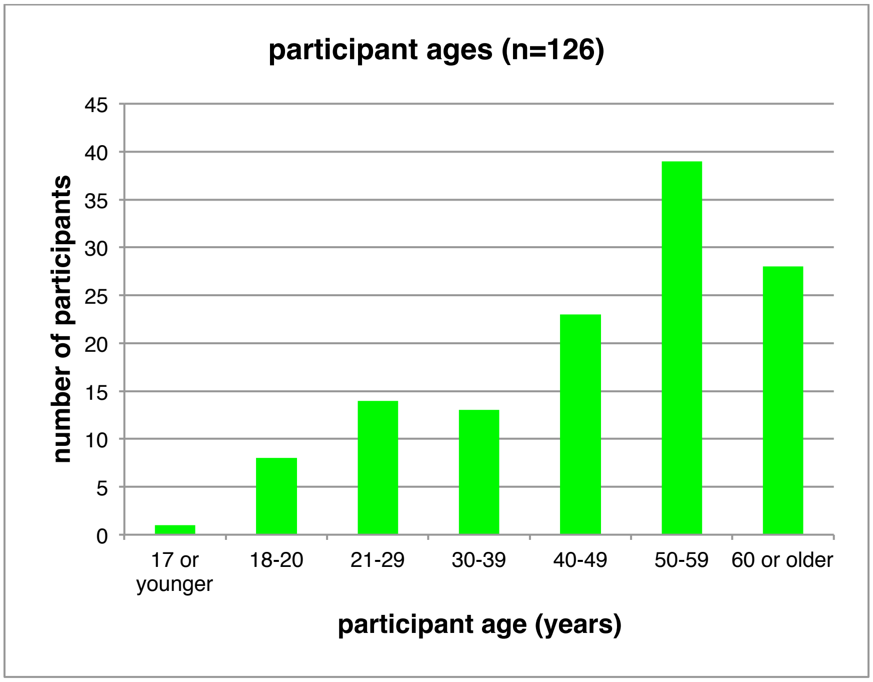


Figure 87: Participant Ages, Online Survey

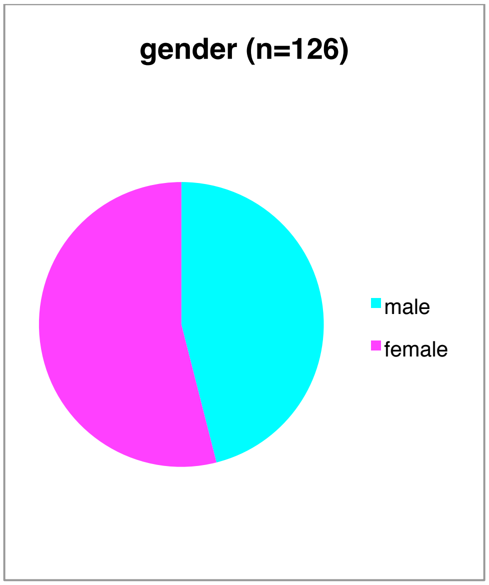


Figure 88: Participant Gender, Online Survey

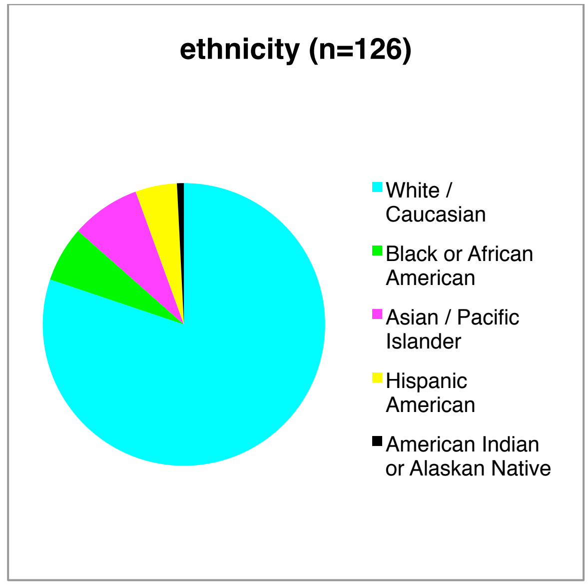


Figure 89: Participant Ethnicity, Online Survey

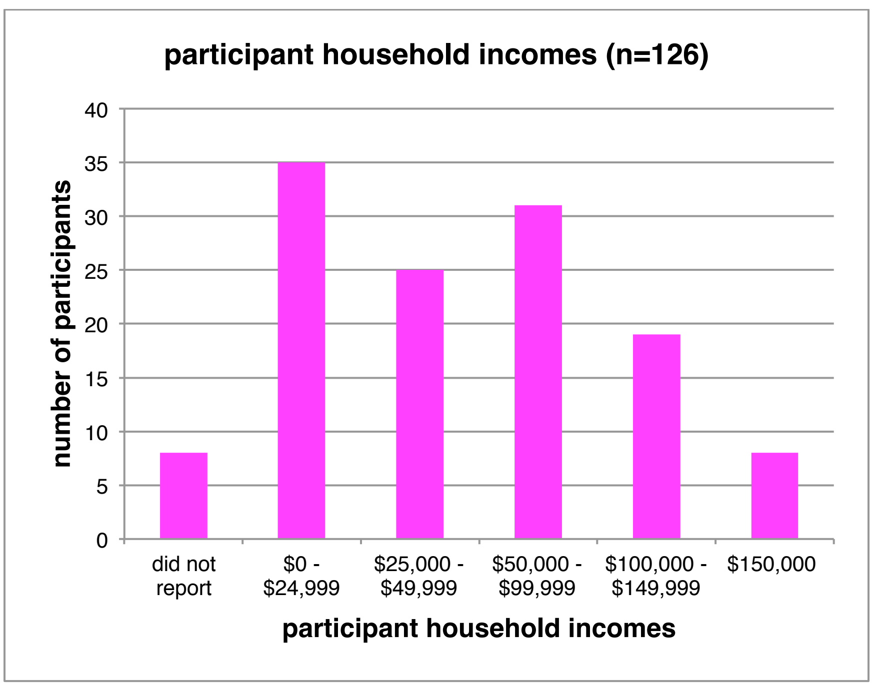


Figure 90: Participant Household Incomes, Online Survey

Appendix D

EDA Peaks per Participant

TABLE 02						
EDA data per participant, individual to normalized						
			(μS, microSiemens)			
	participant	*peaks*	*mean EDA per interval*	*high*	*low*	*normalized EDA*
negative stimulus	A01	**0**	2.55270136	3.686	1.297	0.525617982
	A02	**0**	0.20452944	0.481	0.1	0.274355486
	A03	**0**	0.32557059	2.195	0.2	0.062942652
	A06	**0**	0.93809816	2.458	0.35	0.278983947
	A07	**0**	0.52920982	2.259	0.1	0.198800287
	A08	**0**	2.26098374	5.655	0.8	0.30092353
	A09	**0**	0.72883784	2.732	0.15	0.224181967
	A10	**0**	0.22692667	1.704	0.2	0.017903371
	A11	**0**	0.88701818	2.059	0.05	0.416634236
	A12	**0**	0.076	7.705	0.01	0.008576998
	A13	**0**	0.52625824	0.543	0.15	0.957400102
	A14	**0**	2.09887831	2.269	0.6	0.898069688
	A15	**0**	0.87176748	1.602	0.25	0.459887189
	A17	**0**	1.41470611	3.798	1	0.148215193
	A18	**0**	0.39966489	1.631	0.2	0.139528225
	A20	**0**	0.06305387	0.189	0.05	0.093912734
	A21	**0**	0.51734868	0.793	0.3	0.440869533
	A22	**0**	0.88202174	1.729	0.3	0.40729303
positive stimulus	B01	**0**	1.931705	2.594	0.45	0.69109375
	B02	**0**	0.50784015	8.768	0.5	0.000948252
	B03	**0**	0.38453061	2.93	0.2	0.06759363
	B04	**0**	0.04735789	1.821	0.01	0.020628321
	B05	**2**	0.39633526	1.357	0.05	0.264984897
	B06	**0**	0.10944118	8.107	0.01	0.012281237
	B07	**0**	0.29628691	0.948	0.1	0.231470413
	B10	**0**	0.19098698	3.208	0.15	0.013403198
	B11	**0**	2.80190556	6.04	0.5	0.415506419
	B12	**0**	1.2496	5.265	0.1	0.222575024
	B13	**91**	8.00877128	9.406	1.25	0.828687013
	B14	**0**	2.48666016	2.999	0.55	0.790796309
	B15	**0**	1.25849787	3.151	0.6	0.25813323
	B16	**48**	5.6395625	6.429	1.4	0.843022967
	B17	**0**	2.57115301	5.106	2.1	0.156737528
	B18	**0**	0.05188278	1.564	0.01	0.026951596
	B19	**0**	2.12407818	4.187	0.5	0.440487708
	B20	**17**	4.91266977	5.855	2	0.755556361

Table 2: EDA Peaks Per Participant

Appendix E

Eye Tracking Metrics

TABLE 03			
Participant Eye Tracking Data			
Participant	**TTFF (s)**	**TFD (s)**	**FC (s)**
A01	2.13	2.47	9
A02	1.13	0.5	2
A03	8.7	0.23	1
A06	9.1	0.47	2
A07	0.2	1.9	5
A08	0.07	10.53	23
A09	0.03	3.77	8
A10	2.3	0.43	2
A11	3.63	5.2	16
A12	0.33	1.07	4
A13	2.7	0.6	3
A14	3.37	0.83	3
A15	5.93	13.8	32
A17	0.27	2.4	8
A18	7.87	0.67	3
A20	0.83	4.23	15
A21	2.5	4.37	11
A22	0.07	2.13	8
mean	2.842222222	3.088888889	8.611111111
B01	3.37	1	3
B02	0.03	2.93	7
B03	2.83	4.1	9
B04	12.63	0.73	1
B05	0.6	0.3	2
B06	0.07	10.57	29
B07	0.1	6.1	15
B10	6.63	9.27	26
B11	4.6	2.37	6
B12	4.77	4.07	9
B13	0.03	6.7	19
B14	0.03	2.87	8
B15	0.3	2.2	7
B16	4.63	2.9	7
B17	0.97	0.6	2
B18	0.5	2.53	6
B19	0.03	4.47	15
B20	1.63	1.6	5
mean	2.430555556	3.628333333	9.777777778

Table 3: Eye Tracking Metrics

Appendix F

ANOVA Tables

Electrodermal Activity

TABLE 04						
Analysis of Variance (One-Way): Normalized Mean EDA per Stimulus Interval						
Summary						
Groups	*Sample Size*	*Sum*	*Mean*	*Variance*		
EDA negative	18	5.8541	0.32523	0.07229		
EDA positive	18	6.04086	0.3356	0.09889		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	0.00097	**1**	0.00097	0.01132	**0.91589**	5.95921
Within Groups	2.91006	34	0.08559			
Total	2.91103	35				

Table 4: ANOVA, EDA

Eye Tracking Metrics

TABLE 05						
Analysis of Variance (One-Way): Time to First Fixation						
Summary						
Groups	*Sample Size*	*Sum*	*Mean*	*Variance*		
TTFF negative	18	51.16	2.84222	9.42898		
TTFF positive	18	43.75	2.43056	10.95111		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	1.52523	**1**	1.52523	0.14968	**0.70125**	5.95921
Within Groups	346.46161	34	10.19005			
Total	347.98683	35				

Table 5: ANOVA, TTFF

TABLE 06
Analysis of Variance (One-Way): Total Fixation Duration
Summary

Groups	Sample Size	Sum	Mean	Variance		
TFD negative	18	55.6	3.08889	13.59886		
TFD positive	18	65.31	3.62833	8.3988		
ANOVA						
Source of Variation	SS	df	MS	F	p-level	F crit
Between Groups	2.619	**1**	2.619	0.23812	**0.6287**	5.95921
Within Groups	373.96023	34	10.99883			
Total	376.57923	35				

Table 6: ANOVA, TFD

TABLE 07						
Analysis of Variance (One-Way): Fixation Count						
Summary						
Groups	Sample Size	Sum	Mean	Variance		
FC negative	18	155	8.61111	69.07516		
FC positive	18	176	9.77778	64.4183		
ANOVA						
Source of Variation	SS	df	MS	F	p-level	F crit
Between Groups	12.25	**1**	12.25	0.18353	**0.67106**	5.95921
Within Groups	2,269.38889	34	66.74673			
Total	2,281.63889	35				

Table 7: ANOVA, FC

Self-Assessment Manikin, Part 1: (n=36)

TABLE 08						
Analysis of Variance (One-Way): Happy to Sad (n=36)						
Summary						
Groups	Sample Size	Sum	Mean	Variance		
happy to sad: positive stimulus	36	131	3.63889	2.98016		
happy to sad: negative stimulus	36	277	7.69444	4.10397		
ANOVA						
Source of Variation	SS	df	MS	F	p-level	F crit
Between Groups	296.05556	**1**	296.05556	83.58279	**1.44995E-13**	5.66824
Within Groups	247.94444	70	3.54206			
Total	376.57923	71				

Table 8: ANOVA, Happy to Sad, Thesis Study

TABLE 09						
Analysis of Variance (One-Way): Stimulated to Bored (n=36)						
Summary						
Groups	*Sample Size*	*Sum*	*Mean*	*Variance*		
stimulated to bored: positive stimulus	36	179	4.97222	3.22778		
stimulated to bored: negative stimulus	36	96	2.66667	3.42857		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	95.68056	**1**	95.68056	28.74866	**0**	5.66824
Within Groups	232.97222	70	3.32817			
Total	328.65278	71				

Table 9: ANOVA, Stimulated to Bored, Thesis Study

TABLE 10						
Analysis of Variance (One-Way): Being Controlled to In-Control (n=36)						
Summary						
Groups	*Sample Size*	*Sum*	*Mean*	*Variance*		
being controlled to in control: positive stimulus	36	177	4.91667	3.10714		
being controlled to in control: negative stimulus	36	192	5.33333	8.05714		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	3.125	**1**	3.125	0.55982	**0.45684**	5.66824
Within Groups	390.75	70	5.58214			
Total	393.875	71				

Table 10: ANOVA, Being Controlled to In-Control, Thesis Study

Self-Assessment Manikin, Part 2: (n=126)

TABLE 11						
Analysis of Variance (One-Way): Happy to Sad (n=126)						
Summary						
Groups	*Sample Size*	*Sum*	*Mean*	*Variance*		
happy to sad: positive	126	495	3.92857	3.36286		

stimulus						
happy to sad: negative stimulus	126	935	7.42063	5.33365		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	768.25397	**1**	768.25397	176.68103	**0.E+0**	5.48195
Within Groups	1,087.06349	250	4.34825			
Total	1,855.31746	251				

Table 11: ANOVA, Happy to Sad, Online Survey

TABLE 12						
Analysis of Variance (One-Way): Stimulated to Bored (n=126)						
Summary						
Groups	*Sample Size*	*Sum*	*Mean*	*Variance*		
stimulated to bored: positive stimulus	126	672	5.33333	4.208		
stimulated to bored: negative stimulus	126	528	4.19048	8.13943		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	82.28571	**1**	82.28571	13.3284	**0.00032**	5.48195
Within Groups	1,543.42857	250	6.17371			
Total	1,625.71429	251				

Table 12: ANOVA, Stimulated to Bored, Online Survey

TABLE 13						
Analysis of Variance (One-Way): Being Controlled to In-Control (n=126)						
Summary						
Groups	*Sample Size*	*Sum*	*Mean*	*Variance*		
being controlled to in control: positive stimulus	126	621	4.92857	2.91486		
being controlled to in control: negative stimulus	126	637	5.05556	8.48489		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	1.01587	**1**	1.01587	0.17823	**0.67326**	5.48195
Within Groups	1,424.96825	250	5.69987			
Total	1,425.98413	251				

Table 13: ANOVA, Being Controlled to In-Control, Online Survey

Design Keywords, Part 1: (n=36)

TABLE 14						
Analysis of Variance (One-Way): Design Keywords (n=36)						
Summary						
Groups	*Sample Size*	*Sum*	*Mean*	*Variance*		
positive keywords	360	1,475	4.09722	0.65069		
negative keywords	216	377	1.74537	0.59068		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	746.71296	**1**	746.71296	1,188.63573	**0.E+0**	5.44224
Within Groups	360.59259	574	0.62821			
Total	1,107.30556	575				

Table 14: ANOVA, Design Keywords, Thesis Study

Design Keywords, Part 2: (n=126)

TABLE 15						
Analysis of Variance (One-Way): Design Keywords (n=126)						
Summary						
Groups	*Sample Size*	*Sum*	*Mean*	*Variance*		
positive keywords	1260	4,936	3.91746	0.75887		
negative keywords	756	1,405	1.85847	0.92431		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*p-level*	*F crit*
Between Groups	2,003.14448	**1**	2,003.14448	2,440.21173	**0.E+0**	5.42052
Within Groups	1,653.27169	2014	0.82089			
Total	3,656.41617	2015				

Table 15: ANOVA, Design Keywords, Online Study

Appendix G

Pilot Study Survey

1. What is your ID number? _______

2. Which category below includes your age?
   - ◯ 17 or younger
   - ◯ 18-20
   - ◯ 21-29
   - ◯ 30-39
   - ◯ 40-49
   - ◯ 50-59
   - ◯ 60 or older

3. What is your gender?
   - ◯ Female
   - ◯ Male

4. What is your approximate average household income?
   - ◯ $0-$24,999
   - ◯ $25,000-$49,999
   - ◯ $50,000-$74,999
   - ◯ $75,000-$99,999
   - ◯ $100,000-$124,999
   - ◯ $125,000-$149,999
   - ◯ $150,000-$174,999
   - ◯ $175,000-$199,999
   - ◯ $200,000 and up

5. Are you White, Black or African-American, American Indian or Alaskan Native, Asian, Native Hawaiian or other Pacific islander, or some other race?
   - ◯ White
   - ◯ Black or African-American
   - ◯ American Indian or Alaskan Native

- Asian
- Native Hawaiian or other Pacific Islander
- From multiple races
- Other (please specify)

6. Do you wear any of the following?
   - Glasses
   - Contacts

7. Do you have any of the following?
   - cataracts
   - glaucoma
   - eye implants
   - permanent dilation

8. How many people currently live in your household? ________

9. What is the highest level of school you have completed or the highest degree you have received?
   - Less than high school degree
   - High school degree or equivalent (e.g., GED)
   - Some college but no degree
   - Associate degree
   - Bachelor degree
   - Graduate degree

10. Are you now married, widowed, divorced, separated, or never married?
    - Married
    - Widowed
    - Divorced
    - Separated
    - Never married

11. Are you the primary shopper for your household?
    - Yes
    - No

12. How many children age 17 or younger live in your household? ________

13. Now, please press the circle button ONCE on your LEFT wrist sensor. After pressing the sensor button once, please select "go" to continue.
    - ○ GO: I've pressed my left wrist sensor button!

14. Please describe in detail, in 3-5 complete sentences, your personal feelings towards this package design? How did this make you feel?

____

____

____

____

____

____

15. How would you describe the emotional quality of the package design pictured above?
    I think the package design is:
    - ○ Extremely negative
    - ○ Negative
    - ○ Neutral
    - ○ Positive
    - ○ Extremely positive

16. Do you believe the following words or phrases to be negative, neutral, or positive and to what degree?

	Extremely negative	Negative	Neutral	Positive	Extremely positive
Sustainable	O	O	O	O	O
All-natural	O	O	O	O	O
Sunshine	O	O	O	O	O
Happy	O	O	O	O	O
Healthy	O	O	O	O	O
Recycled	O	O	O	O	O
Packaging	O	O	O	O	O
Smiling	O	O	O	O	O
Hazardous	O	O	O	O	O
Warning	O	O	O	O	O
Tetrachloroethane	O	O	O	O	O
Harmful	O	O	O	O	O
Dangerous	O	O	O	O	O
Black	O	O	O	O	O
Green	O	O	O	O	O
Yellow	O	O	O	O	O
Blue	O	O	O	O	O

17. How natural did your interactions with the environment seem?
   - not at all
   - very little
   - somewhat
   - to a great extent

18. How much did the visual aspects of the environment involve you?
   - not at all
   - very little
   - somewhat
   - to a great extent

19. How much did your experiences in the virtual environment seem consistent with your real world experiences?
   - not at all

- very little
- somewhat
- to a great extent

20. How completely were you able to actively survey or search the environment using vision?
    - not at all
    - very little
    - somewhat
    - to a great extent

21. How much did the control devices (wrist sensor and glasses) interfere with the performance of assigned tasks or with other activities?
    - not at all
    - very little
    - somewhat
    - to a great extent

22. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?
    - not at all
    - very little
    - somewhat
    - to a great extent

23. Were you involved in the experimental task to the extent that you lost track of time?
    - not at all
    - very little
    - somewhat
    - to a great extent

24. How easily did you adjust to the wrist sensors and glasses worn during the shopping experience?
    - not at all
    - very little

- somewhat
- to a great extent

25. Did you feel that CUshop was a realistic shopping experience?
    - not at all
    - very little
    - somewhat
    - to a great extent

26. What could have been done better?

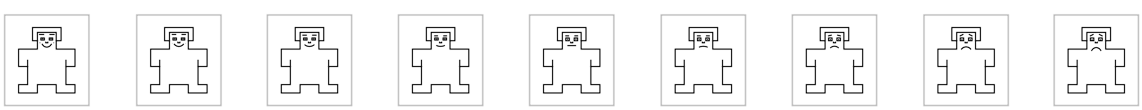
Appendix H

Full Scale Study Survey

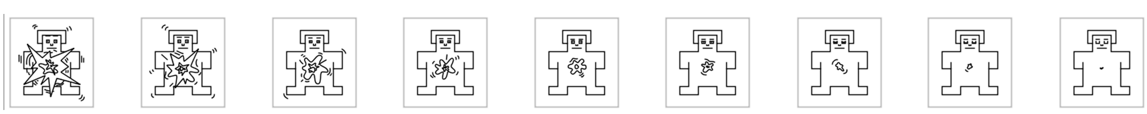
1. Please enter your respondent code here: ________

Please view this package design. How does the design make you feel?
DON'T rate the item. Rate YOUR FEELINGS. There are three rows of emotional characters below. Answer by choosing one emotional character per row.

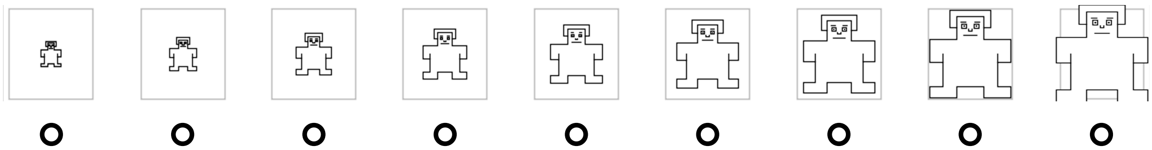
2. Rate your feelings towards the design on a range from happy to sad.



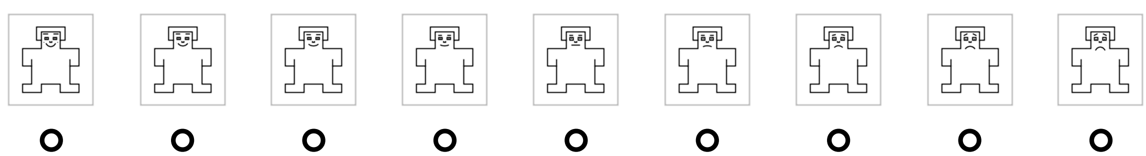
3. Rate your feelings towards the design on a range from stimulated or involved to calm or bored.



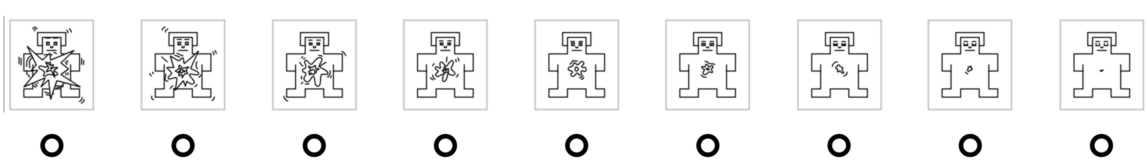
4. Rate your feelings towards the design on a range from being controlled or cared-for to in-control or dominant.



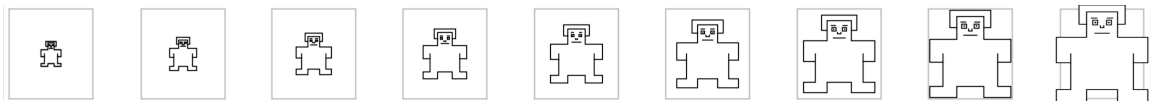
5. Rate your feelings towards the design on a range from happy to sad.



6. Rate your feelings towards the design on a range from stimulated or involved to calm or bored.

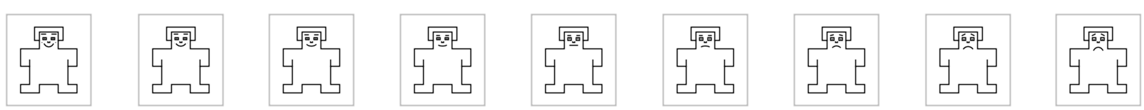


7. Rate your feelings towards the design on a range from being controlled or cared-for to in-control or dominant.



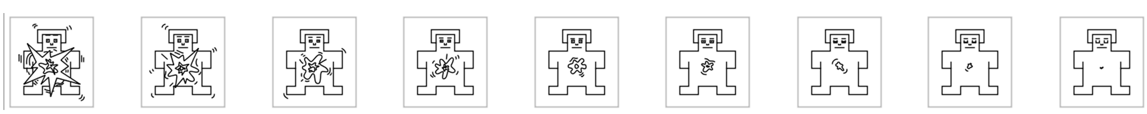
○ ○ ○ ○ ○ ○ ○ ○ ○

8. Rate your feelings towards the design on a range from happy to sad.



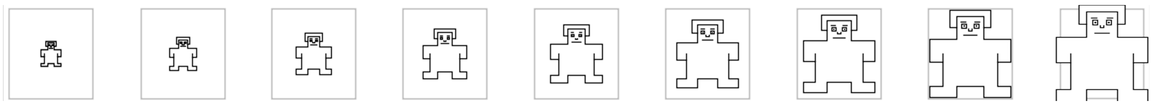
○ ○ ○ ○ ○ ○ ○ ○ ○

9. Rate your feelings towards the design on a range from stimulated or involved to calm or bored.



○ ○ ○ ○ ○ ○ ○ ○ ○

10. Rate your feelings towards the design on a range from being controlled or cared-for to in-control or dominant.



○ ○ ○ ○ ○ ○ ○ ○ ○

---

11. Do you believe the following words or phrases to be negative, neutral, or positive and to what degree?

	Extremely negative	Negative	Neutral	Positive	Extremely positive
Sustainable	○	○	○	○	○
All-natural	○	○	○	○	○
Sunshine	○	○	○	○	○
Happy	○	○	○	○	○
Healthy	○	○	○	○	○
Recycled	○	○	○	○	○
Packaging	○	○	○	○	○
Smiling	○	○	○	○	○
Hazardous	○	○	○	○	○
Warning	○	○	○	○	○
Tetrachloroethane	○	○	○	○	○
Harmful	○	○	○	○	○
Dangerous	○	○	○	○	○
Black	○	○	○	○	○
Green	○	○	○	○	○
Yellow	○	○	○	○	○
Blue	○	○	○	○	○

12. I prefer one-on-one conversations to group activities.
    ○ True
    ○ False

13. I often prefer to express myself in writing.
    ○ True
    ○ False

14. I enjoy solitude.
   - True
   - False

15. I seem to care about wealth, fame, and status less than my peers.
   - True
   - False

16. People tell me that I'm a good listener.
   - True
   - False

17. I'm not a big risk-taker.
   - True
   - False

18. I enjoy work that allows me to "dive in" with few interruptions.
   - True
   - False

19. I like to celebrate birthdays on a small scale, with only one or two close friends or family members.
   - True
   - False

20. People describe me as "soft-spoken" or "mellow."
   - True
   - False

21. I prefer not to show or discuss my work with others until it's finished.
   - True
   - False

22. I tend to think before I speak.
   - True
   - False

23. I often let calls go through voice-mail.
- True
- False

24. I tend to get clammy or sweaty palms when I am nervous.
- True
- False

25. Which category below includes your age?
- 17 or younger
- 18-20
- 21-29
- 30-39
- 40-49
- 50-59
- 60 or older

26. What is your gender?
- Female
- Male

27. Do you wear any of the following?
- Glasses
- Contacts

28. Do you have any of the following?
- Cataracts
- Glaucoma
- Eye Implants
- Permanent Dilation

29. Which race/ethnicity best describes you? (Please choose only one.)
- American Indian or Alaskan Native
- Asian / Pacific Islander
- Black or African American
- Hispanic American

- White / Caucasian

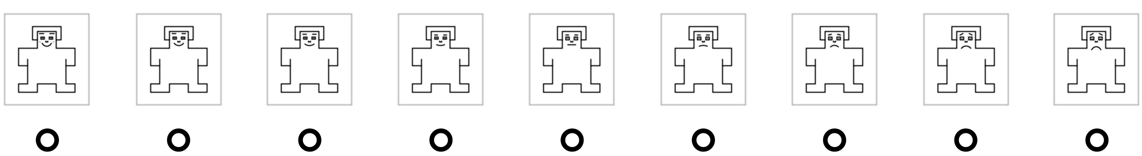
30. What is your approximate average household income?
    - $0-$24,999
    - $25,000-$49,999
    - $50,000-$74,999
    - $75,000-$99,999
    - $100,000-$124,999
    - $125,000-$149,999
    - $150,000-$174,999
    - $175,000-$199,999
    - $200,000 and up

Appendix I

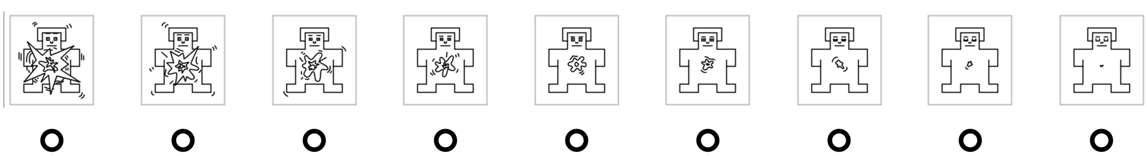
Online Survey Monkey Survey

Please view this package design. How does the design make you feel?
DON'T rate the item. Rate YOUR FEELINGS. There are three rows of emotional characters below. Answer by choosing one emotional character per row.

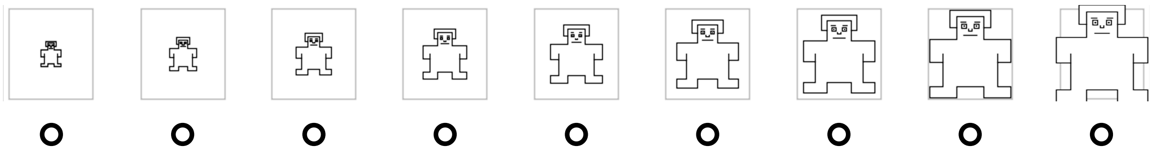
1. Rate your feelings towards the design on a range from happy to sad.



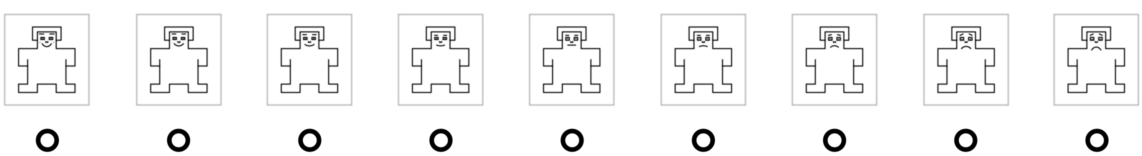
2. Rate your feelings towards the design on a range from stimulated or involved to calm or bored.



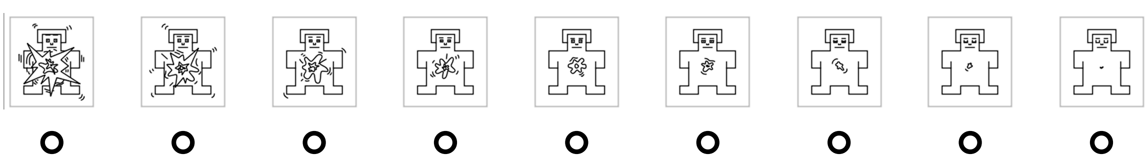
3. Rate your feelings towards the design on a range from being controlled or cared-for to in-control or dominant.



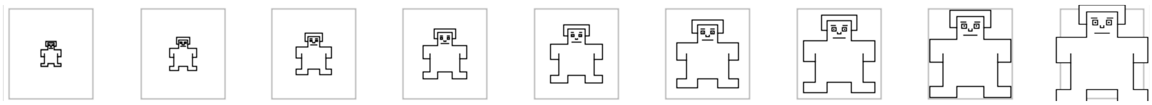
4. Rate your feelings towards the design on a range from happy to sad.



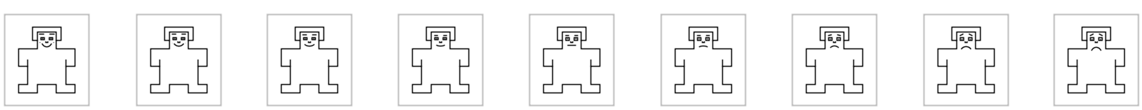
5. Rate your feelings towards the design on a range from stimulated or involved to calm or bored.



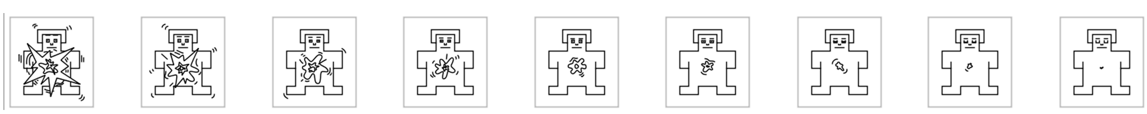
6. Rate your feelings towards the design on a range from being controlled or cared-for to in-control or dominant.



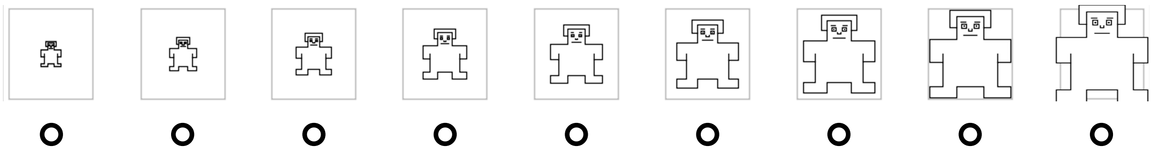
7. Rate your feelings towards the design on a range from happy to sad.



8. Rate your feelings towards the design on a range from stimulated or involved to calm or bored.



9. Rate your feelings towards the design on a range from being controlled or cared-for to in-control or dominant.



10. Do you believe the following words or phrases to be negative, neutral, or positive and to what degree?

	Extremely negative	Negative	Neutral	Positive	Extremely positive
Sustainable	○	○	○	○	○
All-natural	○	○	○	○	○
Sunshine	○	○	○	○	○
Happy	○	○	○	○	○
Healthy	○	○	○	○	○
Recycled	○	○	○	○	○
Packaging	○	○	○	○	○
Smiling	○	○	○	○	○
Hazardous	○	○	○	○	○
Warning	○	○	○	○	○
Tetrachloroethane	○	○	○	○	○
Harmful	○	○	○	○	○
Dangerous	○	○	○	○	○
Black	○	○	○	○	○
Green	○	○	○	○	○
Yellow	○	○	○	○	○
Blue	○	○	○	○	○

11. Which category below includes your age?
   - ○ 17 or younger
   - ○ 18-20
   - ○ 21-29
   - ○ 30-39
   - ○ 40-49

- 50-59
- 60 or older

12. What is your gender?
    - Female
    - Male

13. Do you wear any of the following?
    - Glasses
    - Contacts

14. Do you have any of the following?
    - Cataracts
    - Glaucoma
    - Eye Implants
    - Permanent Dilation

15. Which race/ethnicity best describes you? (Please choose only one.)
    - American Indian or Alaskan Native
    - Asian / Pacific Islander
    - Black or African American
    - Hispanic American
    - White / Caucasian

16. What is your approximate average household income?
    - $0-$24,999
    - $25,000-$49,999
    - $50,000-$74,999
    - $75,000-$99,999
    - $100,000-$124,999
    - $125,000-$149,999
    - $150,000-$174,999
    - $175,000-$199,999
    - $200,000 and up

REFERENCES

Abbas, N., Kumar, D., & Mclachlan, N. (2005). The Psychological and Physiological Effects of Light and Colour on Space Users. *Engineering in Medicine and Biology Society* (p. 2006).

Affectiva. (2013). Emotion Data. No Strings Attached. *Affectiva, Inc.* Retrieved from http://www.affectiva.com/q-sensor/features/

Ahn, H., & Picard, R. (2006). Affective Cognitive Learning and Decision Making: The Role of Emotions. *In The 18th European Meeting on Cybernetics and Systems Research*. Retrieved from http://www.media.mit.edu/affect/pdfs/06.ahn-picard-emcsr.pdf

Andreassi, J. (2000). *Psychophysiology: Human Behavior & Physiological Response*. (L. Erlbaum, Ed.) (4th ed., p. 458). Abingdon, Oxford, UK: Psychology Press.

Bagozzi, R., Gopinath, M., & Nyer, P. (1999). The Role of Emotions in Marketing. *Journal of the Academy of Marketing Science*, *27*(2), 184–206. doi:10.1177/0092070399272005

Belch, M., Holgerson, B., Belch, G., & Koppman, J. (1982). Psychophysiological and Cognitive Responses to Sex in Advertising. *Advances in Consumer Research*, *9*(1), 424–428.

Blakeslee, S. (2004, October 21). Is There A “Buy Button” In the Brain? *The International Herald Tribune*, pp. 1–2. New York, New York, USA.

Boucsein, W, & Hoffmann, G. (1979). A Direct Comparison of the Skin Conductance and Skin Resistance Methods. *Psychophysiology*, *16*(1), 66–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/758630

Boucsein, Wolfram. (1992). *Electrodermal Activity* (2nd ed.). London: Springer Science + Business Media, LLC.

Bradley, M., & Lang, P. J. (1994). Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior and Experimental Psychiatry*, *25*(I), 49–54.

Bradley, M. M., Miccoli, L., Escrig, M. a, & Lang, P. J. (2008). The Pupil As a Measure of Emotional Arousal and Autonomic Activation. *Psychophysiology*, *45*(4), 602–7. doi:10.1111/j.1469-8986.2008.00654.x

Cain, S. (2013). *Quiet: The Power of Introverts in a World That Can't Stop Talking* (1st ed., pp. 1–368). New York, New York, USA: Broadway.

Carbon, C.-C. (2010). The Cycle of Preference: Long-Term Dynamics of Aesthetic Appreciation. *Acta Psychologica*, *134*(2), 233–44. doi:10.1016/j.actpsy.2010.02.004

Chamberlain, L., & Broderick, A. (2007). The Application of Physiological Observation Methods to Emotion Research. *Qualitative Market Research: An International Journal*, *10*(2), 199–216. doi:10.1108/13522750710740853

Chandon, P. (2002). Do We Know What We Look at? An Eye-Tracking Study of Visual Attention and Memory for Brands at the Point of Purchase. *Journal of Consumer Research*, (60), 1–41.

Chew, Y. C., Mappus, C., & Jackson, M. M. (2010). BCI and Creativity. *CHI* (pp. 1–4). Atlanta, GA, USA: CHI.

Coan, J. A., & Allen, J. J. B. (2007). *Handbook of Emotion Elicitation and Assessment*. (J. A. Coan & J. J. Allen, Eds.) (1st ed., pp. 29–46). New York, New York, USA: Oxford University Press.

Consoli, D. (2009). Emotions that Influence Purchase Decisions and their Electronic Processing. *Annales Universitatis Apulensis Series Oeconomica*, *11*(2), 996–1008. Retrieved from http://www.oeconomica.uab.ro/upload/lucrari/1120092/45.pdf

Crowley, A. (1993). The Two-Dimensional Impact of Color on Shopping. *Marketing letters*, *4*(1), 59–69. Retrieved from http://www.springerlink.com/index/U823VPL23K51K321.pdf

Davidson, R. J. (1995). Cerebral Asymmetry, Emotion, And Affective Style. *Brain Asymmetry, emotion and affective style* (pp. 361–387). Cambridge, MA: MIT Press.

Davidson, R. J., & Tomarken, A. J. (1989). Laterality and Emotion: An Electrophysiological Approach. *Handbook of Neuropsychology* (Vol. 3, pp. 419–441). Amsterdam: Elsevier Science Publishers B.V.

Dawson, M., Schell, A. M., & Filion, D. L. (2007). *The Electrodermal System*. *Handbook of ...* (3rd ed., pp. 159–181). Cambridge University Press.

Detenber, B., Simons, R., & Bennett, G. (1998). Roll ' em!: The Effects of Picture Motion on Emotional Responses. *Journal of Broadcasting & Electronic Media*, *42*(1), 113–127.

Dickson, P., & Sawyer, A. (1990). The Price of Knowledge and Search of Supermarket Shoppers. *The Journal of Marketing*, *54*(3), 42–53. Retrieved from http://www.jstor.org/stable/10.2307/1251815

Duchowski, A. (2007). *Eye Tracking Methodology: Theory and Practice* (2nd ed.). New York, New York, USA: Springer.

Ferber, R. (1977). Can Consumer Be Interdisciplinary? *Journal of Consumer Research, Inc.*, *4*(3), 189–192.

Folch-Lyon, E., & Trost, J. (1981). Conducting Focus Group Sessions. *Studies in Family Planning*, *12*(12), 443–449. Retrieved from http://www.jstor.org/stable/10.2307/1965656

George, J. (2005). On Paper, A World of Opportunity. *Packaging World*. Retrieved January 12, 2013, from http://www.packworld.com/package-design/structural/paper-world-opportunity

Girling, R. (2012). *21 Century Design: Shaping Behavior for Preferable Outcomes* (pp. 1–12). Seattle, WA, USA.

Gofman, A., & Moskowitz, H. (2009). Extending Rule Developing Experimentation to Perception of Food Packages with Eye Tracking. *The Open Food Science Journal*, *3*, 66–78.

Goldinger, S. D., & Papesh, M. H. (2012). Pupil Dilation Reflects the Creation and Retrieval of Memories. *Current Directions in Psychological Science*, *21*(2), 90–95. doi:10.1177/0963721412436811

Grabe, M. E., Zhou, S., Lang, A., & Bolls, P. D. (2010). Packaging Television News: The Effects of Tabloid on Information Processing and Evaluative Responses. *Journal of Broadcasting & Electronic Media*, *44*(4), 581–598.

Granholm, E., & Steinhauer, S. (2004). Pupillometric Measures of Cognitive and Emotional Processes. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *52*(1), 1–6. doi:10.1016/j.ijpsycho.2003.12.001

Grifantini, K. (2011, March 3). Wearable Sensor Reveals What Overwhelms You. *MIT Technology Review*, pp. 1–2. Cambridge, MA. Retrieved from http://www.technologyreview.com/news/423185/wearable-sensor-reveals-what-overwhelms-you/?mod=chfeatured

Groeppel-Klein, A. (2005). Arousal and Consumer In-Store Behavior. *Brain research bulletin*, *67*(5), 428–37. doi:10.1016/j.brainresbull.2005.06.012

Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. (2006). The Validity of the Stimulated Retrospective Think-Aloud Method as Measured by Eye Tracking. *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, 1253. doi:10.1145/1124772.1124961

Harris, R. (2006). *Measuring Emotions*. *Marketing Magazine* (pp. 18–20).

Heine, S., Lehman, D., Peng, K., & Greenholtz, J. (2002). What's Wrong with Cross-Cultural Comparisons of Subjective Likert Scales?: The Reference-Group Effect. *Journal of Personality and Social Psychology*, *82*(6), 903–918. doi:10.1037//0022-3514.82.6.903

Hensel, J. S. (1970). *Physiological Measures of Advertising Effectiveness: A Theoretical and Empirical Investigation*. Ohio State University.

Hess, E., & Polt, J. (1960). Pupil Size as Related to Interest Value of Visual Stimuli. *Science*, *132*(3423), 349–350.

Hurley, R. A. (2011). *Persuasive Packaging: An Eye-Tracking Approach to Design*. Clemson University.

Hurley, R. A., Galvarino, J., Thackston, E., Ouzts, A., & Pham, A. (2012). The Effect of Modifying Structure to Display Product vs. Graphical Representation on Packaging. *Packaging Technology and Science*, *online*, 1–8. doi:10.1002/pts

Jacobs, K., & Hustmyer, F. (1974). Effects of Four Psychological Primary Colors on GSR, Heart Rate and Respiration Rate. *Perceptual and motor skills*, *38*(3), 763–766. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4842431

Knufinke, M. (2012). *The Measurement of Arousal by the Means of Electrodermal Activity During an Actually Performed Balance Beam Routine and Observational Learning of the Same Routine*. University of Twente. Retrieved from http://essay.utwente.nl/61606/

Kohan, X. (1968). A Physiological Measure of Commercial Effectiveness. *Journal of Advertising Research*, *8*(4), 46–49.

Kroeber-Riel, W. (1979). Psychobiological Approaches in Consumer Research. *Journal of Consumer Research*, *5*(4), 240–250.

Kroeber-riel, W. (1979). Psychobiological Approaches in Consumer Research. *Journal of Consumer Research*, *5*(4), 240–250.

Krugman, H. (1981). Live, Simultaneous Study of Stimulus, Response Is Physiological Measurement's "Great Virtue". *Marketing News*, *14*(23), 1–19.

Kuchinke, L., Võ, M., Hofmann, M., & Jacobs, A. (2007). Pupillary Responses During Lexical Decisions Vary with Word Frequency but Not Emotional Valence. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, *65*(2), 132–40. doi:10.1016/j.ijpsycho.2007.04.004

LaBarbera, P., & Tucciarone, J. (1995). GSR Reconsidered: A Behavior-Based Approach to Evaluating and Improving the Sales Potency of Advertising. *Journal of Advertising Research*, *35*(5), 33. Retrieved from http://psycnet.apa.org/psycinfo/1996-13185-001

Landwehr, J., McGill, A., & Herrmann, A. (2011). It's Got the Look: The Effect of Friendly and Aggressive "Facial" Expressions on Product Liking and Sales. *Journal of Marketing*, *75*(May), 132–146. Retrieved from http://www.journals.marketingpower.com/doi/pdf/10.1509/jmkg.75.3.132

Lang, A. (1994). Comments on Setting up a Laboratory. In A. Lang (Ed.), *Measuring Psychological Responses to Media Messages* (1st ed., pp. 227–231). Hillsdale, New Jersey, USA: Lawrence Erlbaum Associates, Publishers.

Lang, P., Greenwald, M., Bradley, M., & Hamm, A. (1993). Looking at Pictures: Affective, Aacial, Visceral, and Behavioral Reactions. *Psychophysiology*, *30*(3), 261–73. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8497555

Lang, P. J., Cuthbert, B. N., & Bradley, M. M. (1998). Measuring Emotion in Therapy: Imagery, Activation, and Feeling. *Behavior Therapy*, *29*(4), 655–674. doi:10.1016/S0005-7894(98)80024-5

Larsen, R. J., & Diener, E. (1992). Promises and Problems With the Circumplex Model of Emotion. *Review of Personality and Social Psychology* (13th ed., pp. 25–59). Newbury Park, CA: SAGE Publications.

Latulipe, C., Carroll, E., & Lottridge, D. (2011). Love , Hate , Arousal and Engagement: Exploring Audience Responses to Performing Arts. *CHI 2011: Performing Arts Session* (pp. 1845–1854). Vancouver, BC, Canada. Retrieved from http://dl.acm.org/citation.cfm?id=1979210

Lee, N., Broderick, A., & Chamberlain, L. (2007). What is "Neuromarketing"? A Discussion and Agenda for Future Research. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *63*(2), 199–204. doi:10.1016/j.ijpsycho.2006.03.007

Liao, L., Corsi, A., Lockshin, L., & Chrysochou, P. (2012). Can Packaging Elements Elicit Consumers' Emotional Responses? *41st European Marketing Academy Conference* (pp. 1–7). Lisbon.

Light, L. (1993). At the Center of It All Is the Brand; Promotion Has Bigger Role Than Ads, But "Short-Term Bribes" Are Suicidal. *Advertising Age*, *64*(13), 22.

Lindberg, G. (1988). Advertising Research Techniques: Motivational Response Method. *Agribusiness Marketing Research Conference* (pp. 13–15). San Antonio, Texas.

Lindstrom, M. (2010). *Buyology: Truth and Lies About Why We Buy* (1st ed., pp. 1–272). New York, New York, USA: Crown Business.

Loftus, G. R. (1972). Eye Fixations and Recognition Memory for Pictures. *Cognitive Psychology*, *3*(4), 525–551. doi:10.1016/0010-0285(72)90021-7

Lykken, D., & Rose, R. (1966). Correcting Psychophysiological Measures for Individual Differences in Range. *Psychological Bulletin*, *66*(6), 481–484.

McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., & Montague, P. R. (2004). Neural Correlates of Behavioral Preference for Culturally Familiar Drinks. *Neuron*, *44*(2), 379–87. doi:10.1016/j.neuron.2004.09.019

Morin, C. (2011). Neuromarketing: The New Science of Consumer Behavior. *Society*, *48*(2), 131–135. doi:10.1007/s12115-010-9408-1

Morris, J. (1995). Observations : SAM : The Self-Assessment Manikin An Efficient Cross-Cultural Measurement Of Emotional Response. *Journal of Advertising Research*, (December). Retrieved from ftp://212.74.12.3/OpenShare/Emotion Validation Project/RELATED PAPERS/Observations- SAM-The Self-Assessment Manikin-An Efficient Cross-cultural Measurement of Emotional Response..pdf

Myin-Germeys, I., Van Os, J., Schwartz, J., Stone, A., & Delespaul, P. (2001). Emotional Reactivity to Daily Life Stress in Psychosis. *Archives of General Psychiatry*, *58*(12), 1137–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11735842

Nancarrow, C., Wright, L., & Brace, I. (1998). Gaining Competitive Advantage From Packaging and Labelling in Marketing Communications. *British Food Journal*, *100*(2), 110–118. Retrieved from http://www.emeraldinsight.com/journals.htm?articleid=870420&show=abstract

Nevid, J. (1984). Methodological Considerations in the Use of Electroencephalographic Techniques in Advertising Research. *Psychology and Marketing*, *1*(2), 5–19. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/mar.4220010203/abstract

Nisbett, R. E., & Wilson, T. D. (1977). Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review*, *84*(3), 231–259.

Ohme, R., Reykowska, D., Wiener, D., & Choromanska, A. (2009). Analysis of Neurophysiological Reactions to Advertising Stimuli by Means of EEG and Galvanic Skin Response Measures. *Journal of Neuroscience, Psychology, and Economics*, *2*(1), 21–31. doi:10.1037/a0015462

Ohme, R., Reykowska, D., Wiener, D., & Choromanska, A. (2010). Application of Frontal EEG Asymmetry to Advertising Research. *Journal of Economic Psychology*, *31*(5), 785–793. doi:10.1016/j.joep.2010.03.008

Pawle, J., & Cooper, P. (2006). Measuring Emotion-Lovemarks, the Future Beyond Brands. *Journal of Advertising Research*, 38–48. doi:10.2501/S0021849906060053

Penn, D. (2006). Looking for the Emotional Unconscious in Advertising. *International Journal of Market Research*, *48*(5), 515–524.

Petrides, K. V., & Furnham, A. (2000). On the Dimensional Structure of Emotional Intelligence. *Personality and Individual Differences*, *29*(2), 313–320. doi:10.1016/S0191-8869(99)00195-6

Picard, R., & Daily, S. (2005). *Evaluating Affective Interactions: Alternatives to Asking What Users Feel*. *CHI Workshop on Evaluating Affective …* (pp. 1–4). Portland. Retrieved from http://mit.sustech.edu/NR/rdonlyres/Media-Arts-and-Sciences/MAS-630Spring-2008/Readings/picard_chi2005.pdf

Poels, K., & Dewitte, S. (2006). How to Capture the Heart? Reviewing 20 Years of Emotion Measurement in Advertising. *SSRN Electronic Journal*. doi:10.2139/ssrn.944401

Poh, M., Swenson, N., & Picard, R. (2009). Comfortable Sensor Wristband for Ambulatory Assessment of Electrodermal Activity. *1st Biennial Conference of the Society for Ambulatory Assessment* (p. 2009). Greifswald, Germany. Retrieved from http://affect.media.mit.edu/pdfs/09.Poh-Swenson-Picard-SAA.pdf

Poh, M.-Z., Loddenkemper, T., Swenson, N. C., Goyal, S., Madsen, J. R., & Picard, R. W. (2010). Continuous Monitoring of Electrodermal Activity During Epileptic Seizures Using a Wearable Sensor. *32nd Annual International Conference of the IEEE EMBS* (Vol. 2010, pp. 4415–8). Buenos Aires, Argentina. doi:10.1109/IEMBS.2010.5625988

Poh, M.-Z., Swenson, N. C., & Picard, R. W. (2010). A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity. *IEEE transactions on bio-medical engineering*, *57*(5), 1243–52. doi:10.1109/TBME.2009.2038487

Pop, C. M., Radomir, L., Ioana, M. A., & Maria, Z. M. (2009). Neuromarketing-Getting Inside the Customer's Mind. *Annals of Faculty of Economics*, *4*(1), 804–807. Retrieved from http://econpapers.repec.org/article/orajournl/v_3a4_3ay_3a2009_3ai_3a1_3ap_3a804-807.htm

Pradeep, A. K. (2010). *The Buying Brain: Secrets to Selling to the Subconscious Mind* (1st ed.). Hoboken, NJ: Wiley.

Raghubir, P., & Greenleaf, E. (2006). Ratios in Proportion: What Should the Shape of the Package Be? *Journal of Marketing*, *70*(2), 95–107. Retrieved from http://www.jstor.org/stable/10.2307/30162088

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, *124*(3), 372–422.

Rebollar, R., Lidón, I., Serrano, A., Martín, J., & Fernández, M. J. (2012). Influence of Chewing Gum Packaging Design on Consumer Expectation and Willingness to Buy: An Analysis of Functional, Sensory and Experience Attributes. *Food Quality and Preference*, *24*(1), 162–170. doi:10.1016/j.foodqual.2011.10.011

Reimann, M., Zaichkowsky, J., Neuhaus, C., Bender, T., & Weber, B. (2010). Aesthetic Package Design: A Behavioral, Neural, and Psychological Investigation. *Journal of Consumer Psychology*, *20*(4), 431–441. doi:10.1016/j.jcps.2010.06.009

Rettie, R., & Brewer, C. (2000). The Verbal and Visual Components of Package Design. *Journal of Product & Brand Management*, *9*(1), 56–70. doi:10.1108/10610420010316339

Rundh, B. (2009). Packaging Design: Creating Competitive Advantage with Product Packaging. *British Food Journal*, *111*(9), 988–1002. doi:10.1108/00070700910992880

Ruskin, G. (2004). Letter to John McCain: Commercial Alert Asks Senate Commerce Committee to Investigate Neuromarketing. Retrieved from http://www.commercialalert.org/issues/culture/neuromarketing/commercial-alert-asks-senate-commerce-committee-to-investigate-neuromarketing

Shapiro, M. A. (1994). Think-Aloud and Thought-List Procedures in Investigating Mental Processes. In A. Lang (Ed.), *Measuring Psychological Responses to Media Messages* (1st ed., pp. 1–14). Hillsdale, New Jersey, USA: Lawrence Erlbaum Associates, Publishers.

Shiv, B., & Fedorikhin, A. (1999). Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making. *Journal of Consumer Research*, *26*(3), 278–292.

Silayoi, P., & Speece, M. (2004). Packaging and Purchase Decisions: An Exploratory Study on the Impact of Involvement Level and Time Pressure. *British Food Journal*, *106*(8), 607–628. doi:10.1108/00070700410553602

Singh, S. N., & Churchill, G. A. J. (1987). Arousal and Advertising Effectiveness. *Journal of Advertising*, *16*(1), 4–10.

Sørensen, J. (2008). Measuring Emotions in a Consumer Decision-Making Context – Approaching or Avoiding. Aalborg East.

Stewart, D., & Furse, D. (1982). Applying Psychophysiological Measures to Marketing and Advertising Research Problems. *Current Issues and Research in Advertising*, *5*(1), 1. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/01633392.1982.10505319

Sukhvinder, O. (2011). Book Review: The Buying Brain: Secrets to Selling the Unconscious Mind. *Journal of Consumer Marketing*, *28*(2), 162–164.

Thaler, R. H., & Sustein, C. R. (2009). *Nudge: Improving Decisions about Health, Wealth, and Happiness* (1st ed.). New York, New York, USA: Penguin Group.

Tobii. (2010). *Product Description: Tobii Glasses Eye Tracker* (pp. 1–8). Danderyd, Sweden. Retrieved from www.tobii.com

Tonkin, C., & Duchowski, A. (2011). *CUshop: A Simulated Shopping Environment Fostering Consumer-Centric Packaging Design & Testing*. Retrieved from http://dl.acm.org/citation.cfm?id=2337889

Tonkin, C., Ouzts, A., & Duchowski, A. (2011). Eye Tracking Within the Packaging Design Workflow: Interaction with Physical and Virtual Shelves. *Proceedings of the 1st Conference on Novel Gaze-Controlled Applications* (pp. 1–8). New York: ACM. Retrieved from http://www.tobii.com/Global/Analysis/Marketing/Research Paper/Marketing and media/EyeTracking Within the Packaging Design Workflow.pdf

Tversky, B. (1974). Eye Fixations in Prediction of Recognition and Recall. *Memory & Cognition*, *2*(2), 275–278. Retrieved from http://www.springerlink.com/index/B307J26460728L61.pdf

Tversky, B. (2003). Structures Of Mental Spaces: How People Think About Space. *Environment & Behavior*, *35*(1), 66–80. doi:10.1177/0013916502238865

Valentino-DeVires, J. (2010, January 19). MIT Researchers Read Consumers' Faces to Make a Better Taste Test. *The Wall Street Journal*, p. 1. New York, New York, USA.

Van der Laan, L., De Ridder, D., Viergever, M., & Smeets, P. (2012). Appearance Matters: Neural Correlates of Food Choice and Packaging Aesthetics. *PloS ONE*, *7*(7), e41738. doi:10.1371/journal.pone.0041738

Vecchiato, G., Astolfi, L., De Vico Fallani, F., Toppi, J., Aloise, F., Bez, F., Wei, D., et al. (2011). On the Use of EEG or MEG Brain Imaging Tools in Neuromarketing Research. *Computational intelligence and neuroscience*, *2011*, 643489. doi:10.1155/2011/643489

Wahlberg, D. (2004, February 1). Advertisers Probe Brains, Raise Fears. *The Atlanta Journal-Constitution*, p. Q1. Atlanta, GA.

Wansink, B. (1996). Can Package Size Accelerate Usage Volume? *The Journal of Marketing*, *60*(3), 1–14. Retrieved from http://www.jstor.org/stable/10.2307/1251838

Warde, A. (1999). Convenience Food: Space and Timing. *British Food Journal*, *101*(7), 518–527. doi:10.1108/00070709910279018

Westerman, S. J., Sutherland, E. J., Gardner, P. H., Baig, N., Critchley, C., Hickey, C., Mehigan, S., et al. (2013). The Design of Consumer Packaging: Effects of Manipulations of Shape, Orientation, and Alignment of Graphical Forms on Consumers' Assessments. *Food Quality and Preference*, *27*(1), 8–17. doi:10.1016/j.foodqual.2012.05.007

Whang, M. C., Lim, J. S., & Boucsein, W. (2003). Preparing Computers for Affective Communication: A Psychophysiological Concept and Preliminary Results. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(4), 623–634. doi:10.1518/hfes.45.4.623.27095

Wickelmaier, F. (2003). *An Introduction to MDS* (pp. 1–26). Aalborg East, Denmark.

Wierzbicka, A. (1999). *Emotions Across Languages and Cultures: Diverstiy and Universals* (1st ed.). Cambridge, UK: Cambridge University Press.

Witchalls, C. (2004, March 22). Pushing the Buy Button; Companies Are Starting to Turn to Powerful Brain-Scan Technology in Order to Figure Out How We Choose Which Products to Purchase. *Newsweek International*, p. 50.

Witmer, B., & Singer, M. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence*, *7*(3), 225–240. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/105474698565686

Zaccai, G. (2012, October). Why Focus Groups Kill Innovation, from the designer behind Swiffer. *Fast Co. Design, Fast Company*.