Thick Psychophysiology for Empathic Design

by

Elliott Bruce Hedman

B.S., University of Colorado (2008)

M.S., Massachusetts Institute of Technology (2010)

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
In partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2014

Author____

Program in Media Arts and Sciences
May 5, 2014

Certified by____

Rosalind W. Picard
Professor of Media Arts and Sciences
Program in Media Arts and Sciences, MIT
Thesis supervisor

Accepted by____

Pattie Maes
Associate Academic Head
Program in Media Arts and Sciences, MIT

Thick Psychophysiology for Empathic Design

by

Elliott Bruce Hedman

Submitted to the Program in Media Arts and Sciences
School of Architecture and Planning,
On May 2, 2014, in partial fulfilment of the
Requirements for the degree of
Doctor of Philosophy

Abstract

Over the course of six years, I brought ambulatory psychophysiology into a variety of industries as a means of conducting design research. I looked at the stress of children in occupational therapy, the frustration of playing Hasbro board games, the thrill of driving a Google Self Driving Car, the confidence of shopping at Best Buy and Lowes, the excitement of playing LEGO Technic for the first time, the tension of watching one's first symphony, and the anxiety of talking about birth control. Working with stake holders within these settings I developed "Thick Psychophysiology," defined by four characteristics: 1. Psychophysiological data is quantitatively measured, 2. The research answers explorative, open ended questions, 3. The research measures external context, and 4. The research measures internal context. By combining ethnographic methods with psychophysiology, researchers can address the challenges of specificity that ambulatory, explorative research produces. Two case studies of preliminary design research are provided about the LEGO Group and the New World Symphony, showcasing how thick psychophysiology can help uncover customer's unarticulated needs.

Once needs are uncovered, the challenge is how to motivate an organization to address those needs. Traditionally, designers use storytelling as a way to communicate research findings in regards to user experience, which in some cases can be ineffective in creating the needed motivation. The method developed in this thesis contains components designed to help influence organizational change. To test the effect psychophysiological data can have on organizational change, I delivered a survey testing four ways (conditions) of presenting findings: Storytelling (the most common method used by companies such as IDEO), Video-based (adding video to the story), and two conditions using Video and psychophysiology, varying how the physiological insights were presented (narrow vs. broad). Participants in the broad condition were told the skin conductance could mean a variety of things including moving, breathing hard, being stressed, or being excited. We analyzed the results of 143 LEGO employees. Participants in the broad skin conductance condition had a 47% chance of increasing the priority of the proposed initiative, whereas only 9% of participants in the storytelling condition increased the priority of that initiative ($p<0.01$). Post hoc analysis showed that when participants reported an empathic response to the skin conductance, they were even more likely to increase the priority of that broad skin conductance initiative (75%). These results suggest that, when compared to storytelling, presenting psychophysiological data can be a more effective way to communicate customer experience.

Thesis Supervisor: Rosalind W. Picard

Title: Professor of Media Arts and Sciences, Program in Media Arts and Sciences, MIT

Thick Psychophysiology for Empathic Design

By

Elliott Bruce Hedman

Reader:____

Karen S. Quigley
Research Associate Professor of Psychology
Northeastern University

Reader:____

Frances Frei
Professor of Technology and Operations Management
Harvard Business School

Acknowledgements

Often times, when one goes on an unexpected journey, it is the people we meet along the way that affect us the most. I think of my dissertation as a compilation of all the things people from all the corners of the world taught me about psychophysiology, ethnographic research, and design. And with this framework, there are many who I would like to thank for their contribution to this dissertation.

Thanks to my family for all their support. Every January break, upon coming home from the latest load of finals or paper submissions, I would be greeted by the love of my parents and brothers, then would pass out after a home cooked meal - the best safety net, motivation, and inspiration.

Thick psychophysiology started out at the STAR Center in Colorado, ran by Dr. Lucy Jane Miller. Thanks to Dr. Miller, the researchers, and therapists, who were kind enough of to let me into their community over the summer and who worked together with me to figure out how to make these sensors useful. And thanks to all the children and families in the therapy for inviting me into their lives and helping test out the sensors. A special thanks to Holly Adinoff, Alex Bidwell, Megan Michelle, and Meredith Wilson who put in serious hours placing sensors on children and videotaping them for over 70 hours as we started to unwind what it meant to design a better therapy.

Cecilia Weckstrom from the LEGO Group believed in me before I even believed in myself. She went out of her way to bring the MOXO sensors to the LEGO Group and funded the initial research that helped combine the MOXO sensors with design work. It was there where I met Morten Lundholm who pushed me to think like an ethnographer and designer. And as an electrical computer engineer, I was heavily resistant to those ideas; thank goodness Morten was especially persuasive and persistent. Without his help, I would have never been able to embrace design like I do today. Thanks to Kay Davis and Mette Andresen for bringing that initial LEGO work to a place of pride for me.

A special thanks needs to be sent to Dr. Susan Silbey and Dr. Graham Jones who welcomed me into the Anthropology Department at MIT. Both professors took the time and energy to help me understand research methodologies and theories from a sociological perspective. It was their view of the world that added the much-needed “thick” to my psychophysiology.

Tom Chi from GoogleX reminded me that twisting the discipline of psychophysiology is just the beginning. Tom mentored me while working with the Self Driving Car team and more importantly taught me the importance of rapid prototyping and evaluation. Changing my academic timeline to one-day evaluations allowed me to quickly improve the sensor technology and also arrive at deep, meaningful insights more quickly. Tom's advice has lead me to focus just as much on prototyping as I do on sensors in my current work.

Thanks to Diego Rodriguez, Colin Raney, and Joi Ito for bringing me into the IDEO family as well. The IDEO community introduced me to many peers who were passionate about redefining design and transforming companies. It was here I crafted my design research skills and gained the confidence to discover insights that could transform companies. It was at IDEO I learned how design research fits into the bigger picture of design, the importance of interacting with clients, what quality design looks like, and how sensors can help make design even better. An extra shout out goes to the 70 people who made up the "Elliott Makeover Committee at IDEO" who not only helped me finish my dissertation, but made me look good doing it.

A special thanks goes to the individuals within my first consulting projects: Best Buy, the LEGO Group, the New World Symphony, and Lowe's Hardware. Michael Oshea, Bo Thomsen, Tina Sorensen, Craig Hall, and Adriane Cooley all put in many late night hours bringing these sensors into their respective companies. Those individuals and many others believed that psychophysiology could help their company become user-centered, and they took a risk working with a graduate student whose ideas were much greater than where he started. It was with these companies that I was able to hone my different work and prototype methodologies to figure out how sensors might transform a company. Without their trust, I would never have been able to grow these ideas to fruition.

And thanks to all of the mentors who have helped me figure out who I am within the design world. The dinners with Hugh Dubberly helped me understand the history of design and how psychophysiology could play a role. Dr. James Russell warped my mind multiple times around what it means to measure an emotion and why my assumptions about emotions were mostly wrong. Dr. Frances Frei helped me add the "business" dimension to my work. Her unrelenting investment and encouragement kept me seeking out more.

Dr. Karen Quigley helped me bring my psychophysiology work from an idea to a practice. She helped me understand and appreciate the vast amount of research done in psychophysiology and why ambulatory psychophysiology is much harder than just a problem of making sensors wireless. Together we would go through the data I was collecting and discuss how reliable the data was and what could actually be concluded – something the business world never really asked. It was with Dr. Quigley's help that I learned what I could and could not conclude with the MOXO sensors, and correspondingly learned about the need to add ethnographic methods to my work.

I often refer to my adviser, Dr. Rosalind Picard, as my other mom. Six years ago, I started out knowing very little about machine learning, emotions, psychophysiology, or pretty much anything else. But Dr. Picard invested in me – giving me the resources, time, and encouragement to take risks and try things in a very different way than psychophysiologists or designers had before. I was lucky to have an adviser who would allow me to go ahead and spend 3 months in an occupational therapy setting "to see what would happen." Roz introduced me to mentors and future collaborators across multiple disciplines and encouraged me to formalize my ideas about "Thick Psychophysiology." I had a high bar coming into MIT, but Roz kept raising it higher. Our relationship has been amazing – she allowed me to explore and create, while at the same time expecting the best out of me. Roz helped me understand and pursue my passions, and gave me the tools I needed to make meaningful impact. I remember one workshop, after presenting to several vice president level Media Lab Members, Roz pulled me into her office. She told me the presentation was great, but she was really hoping I would hand out pinwheels to everyone next time. And that so perfectly summed up Roz's impact on me – expects the best I can give, but then pushes me to do things beyond what is normally expected. And so if you ever receive a pinwheel from me, that was at the suggestion of Roz. (You can read about the pinwheels in the Conclusion chapter.)

Thanks everyone for supporting me and inspiring me along this journey. It has truly been an adventure, and the relationships I have formed with you have made it all worthwhile.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 SCENARIO

Occupational therapists at the STAR Center in Denver, Colorado wanted to know how they could improve therapy for children with sensory challenges and Autism Spectrum Disorder. With wireless skin conductance sensors (Fletcher et al., 2010), we measured children's skin conductance as they participated in 70 hours of therapy (Hedman et al., 2012). Lab research suggests that changes in skin conductance can occur when a person is stressed, excited, frustrated, thinking hard, breathing hard, or moving hard (Boucsein et al., 2007). One particular interest of therapists was a child's arousal. People with low arousal are described as calm or relaxed, and people with high arousal may have intense anger, stress, or excitement. Researchers describe skin conductance as being only innervated by the sympathetic nervous system, allowing scientists to measure sympathetic nervous system arousal with skin conductance (Critchley, 2002b).

Unlike traditional psychophysiological studies, we did not know how skin conductance or a child's arousal would change in a natural therapy. Take for example Kara's[1] psychophysiological response during one therapy session.

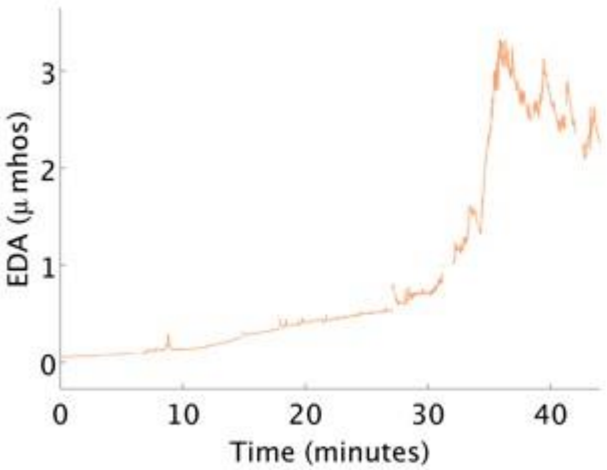


Figure 1.1 Kara's skin conductance without context is difficult to interpret.

[1] Names used in this dissertation are pseudonyms.

Multiple large skin conductance responses occurred during minute 30 to 40, but why? Was Kara anxious? Was she excited? Or perhaps she was moving around? And even if we conclude this is an emotional response, a secondary question becomes what event caused the response? What occurred during therapy that created such a large response for Kara? With only the skin conductance data, we cannot infer much from Kara's reaction.

Data scientists often suggest measuring additional input in order to find correlations between events. For example, I labeled the activities that Kara participated in across time.

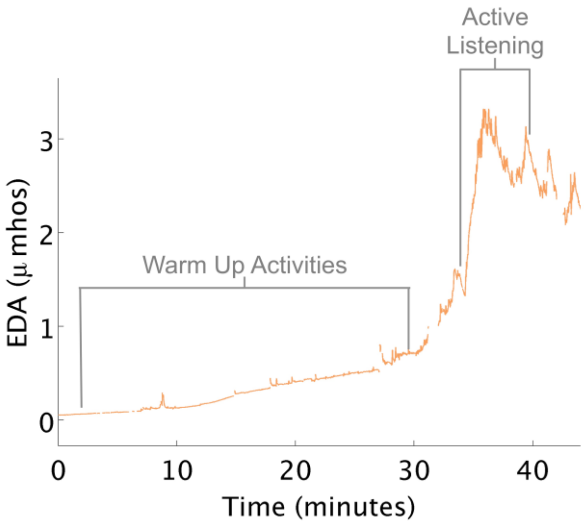


Figure 1.2 By labeling activities, the response becomes clearer.

With this new environmental information, one can tell responses occurred after warm up activities and at the start of Active Listening. Perhaps beginning Active Listening is arousing? But this label fails to describe why a response occurred. Was the child excited or stressed or was this response due to movement? If anxious, what exactly was the child anxious about?

In addition to formal activity metrics, I learned what was occurring during this therapy session, in large part, by accident. When I signed Kara up for the study, I talked to the parents and therapist about Kara – she had Autism Spectrum Disorder and could only speak in one to three word phrases. While Kara was participating in therapy, I watched the sessions from the corner of the room in order to collect the data and videotape the sessions (we needed a way to reference when activities occurred).

From the video and watching the therapy live, I knew the session was early in the morning. Both Kara and I were groggy from waking up so early. The therapist was trying to arouse Kara, so that she could be more active in the therapy. Kara played in a ball pit, jumped on a trampoline, and crawled through a tunnel, with her skin conductance level not changing significantly. Throughout this time, Kara repeatedly said, “Wagon. Wagon.” When the warm up activities finished, Kara rode in a green wagon to Active Listening. While Kara was sitting in this wagon, her skin conductance increased. Kara’s skin conductance continued to climb as she put on headphones to hear her own voice.

Later, I informally showed the therapist Kara’s skin conductance with video, out of curiosity. She pointed out the importance of the wagon to Kara. Kara had too low of arousal during the early morning and nothing seemed to work, but having Kara move to Active Listening seemed to increase that arousal. The therapist joked that Active Listening probably increased Kara’s arousal too much. The therapist told me how Active Listening probably overwhelmed Kara, and having her calmer at the end would have been better. Considering the observations, the therapists report, and background history together, I conclude that Kara’s large skin conductance was likely due to her excitement of riding in a wagon, the reaction to a transition between environments, and the anxiety of having to do Active Listening.

In essence, during the occupational therapy study, I was learning how to do observational research in order to make meaning out of the real-world data I was collecting. I was not the first person to do observational methods. A year later, I learned how anthropologists and sociologists conducted detailed, rich observations to learn about emotion. Take Briggs’s study of emotions with an Eskimo tribe as an example.

1.2 THICK PSYCHOPHYSIOLOGY DEFINED

Briggs lived as an adopted daughter in an igloo for 7 months (1970). In a northern bay of Canada, she interviewed the Utku about their emotions and also observed the tribe’s behavior during emotional events. The Utku experienced anger much differently than Briggs; few outbursts of anger ever occurred. Briggs describes a daughter of the family, Raigili, responding to conflict with her younger sister Saark, while her father Inuttiaq intervenes.

> “As before, Ragili had it [a plastic bag] and Saarak wanted it. This time Inuttiaq intervened. “Give it to Saarak,” he said quietly. Raigili sat immobile with lowered head, while Inuttiaq waited. “Give it to Saarak,” Inuttiaq repeated in the same still voice. Raigli sat, frozen.”

Raigli's behavior of shutting off in times of distress was the norm for her and the other children of the village. By living in a village, interviewing group members, and watching behavior in detail, Briggs learned how anger manifested in the Utku tribe.

My first question, reading Briggs's account, was how did Raigli physiologically respond to these moments of stress? While behaviorally freezing, did her sympathetic nervous system continue to activate? Did shutting down help her decrease her physiological arousal? While Briggs was able to describe the outward behavior of the Utku tribe, as a psychophysiologist, I wanted to know how the tribe's psychophysiological responses differed as well.

"When the distressful events occur to the Utku, how do they physiologically respond?" is an open-ended, exploratory question (Babbie, 2013). When Briggs spent 7 months with the Eskimos, she did not have a pre-established hypothesis that Eskimos would behave differently with anger. She learned about Utku anger through thoughtful, ambulatory observations and grounded theory (Glaser & Strauss, 2009). Exploratory questions can help researchers understand how people physiologically react in day-to-day life. When do psychophysiological responses occur? How do psychophysiology and behavior relate? What methods do people use to regulate or change their psychophysiology? These questions would be particularly hard to test in a laboratory, requiring observations about real-world responses.

Design researchers use exploratory questions as a base of study. Very seldom is there a pre-defined hypothesis to test in design work. Rather, designers hope to understand the current system around a product or service as a whole to help inspire how the system might improve (Dubberly & Evenson, 2008). These exploratory questions are referred to as "diagnostic" in the advertising world (LaBarbera & Tucciarone, 1995). Rather than categorize an ad as good or bad, exploratory questions help describe what is or is not working with an ad. Take for example a study on shopper's skin conductance at an Austrian grocery store. Groeppel-Klein had participants shop in a pre-designated well-designed grocery store, and a poorly-designed grocery store and noted the difference in skin conductance response count (2005). As a designer, I have less interest in validating whether a store is designed well, and I would rather research what works well and does not work well in a space. In design terminology, I am interested in "customer pain points" or the "customer journey." Exploratory questions in grocery shopping might be, "When are shoppers stressed or excited?" or "How could we change the emotional response to change shopping behavior?"

As discussed in Chapter 2, the more exploratory a psychophysiological study is, the more challenging inference becomes. In particular, physiological signals often lose specificity as they are removed from a controlled environment. If participants sit still in a controlled environment with no additional stimuli and watch a movie that is predetermined to make them angry, then any recorded physiological differences between controls are likely due to the anger stimuli (see Wilhelm, 2010 for a counter-argument for why a controlled experiment could be a poor measurement of anger). Whereas, if Regili showed a skin conductance response while sulking in the corner, what could be inferred? Are the skin conductance responses due to anger? Did she change her body position? Is she breathing hard? Is she excited? The list of possible reasons for a response continues indefinitely. The internal and external context of a physiological response is essential to infer the meaning of a response in ambulatory settings.

As anthropologists and sociologists use exploratory, ambulatory research as a standard form of practice, understanding context is a foundation of ethnographic methods. In a seminal piece on methodology, Geertz argues that research needs thick description (1973). He asks how a researcher determines if a wink is a physical reaction or conspiratorial in nature, concluding the researcher must understand the context. A researcher must not just explain the behavior but the context of the behavior as well. Briggs knew that the tears from Regili were due to sadness — not from a harsh snow storm and not from happiness — because she saw Regili lose her toy bag and her father discipline her. Similarly, exploratory psychophysiological questions need thick description. I needed contextual understanding to know that the wagon was important to Kara and was a likely contributor to Kara's physiological response.

In Chapter 3, I describe how to make inferences from exploratory research with a method I title "Thick Psychophysiology." Thick psychophysiology uses observational methods and interviews to gain the thick description of an ambulatory, psychophysiological event. Thick psychophysiology has four defining characteristics:

1. Quantitatively measures a physiological aspect of an individual.
2. Answers an exploratory, ambulatory question.
3. Researches external context: what events lead to a physiological response occurring?

4. Researches internal context: what is the meaning and interpretation of the physiological response?

In Kara's therapy example, the external context includes information like: the therapy session was early in the morning, the child mentioned "wagon" multiple times, and the child had to listen to her own voice. The internal context describes the child's emotional reaction – the therapist reported that the child seemed tired and unresponsive in the morning and appeared overwhelmed and resistive at the end. Kara's experience is one example of thick psychophysiology, with the remainder of the dissertation demonstrating the need and execution of the method in more detail.

1.3 CONTRIBUTIONS

Over the course of my Ph.D. work, I:

1. Addressed explorative, empathic design questions with psychophysiology. I have measured psychophysiological data from people having diverse experiences such as audience members at symphony concerts, drivers in self-driving cars, mothers playing board games, and shoppers choosing new products. In these settings, stakeholders wanted me to help them better understand their customers with psychophysiology as a lens.
2. Combined ethnographic and psychophysiological methods together to allow for specificity in ambulatory settings. In these explorative studies, I often ran the risk of low specificity as people could have had a large range of emotional, cognitive, and physical responses to the open situations tested. Observational methods helped me to identify possible reasons for skin conductance responses and artifacts. I leaned heavily on unstructured, long interviews to help me understand how customers interpreted their reactions. Ethnographic methods helped me understand the context, influence, and internal state of participants, which made inference of explorative psychophysiology feasible.
3. Demonstrated how psychophysiology can be an effective way for design researchers to share knowledge across an organization. From a survey with 143 LEGO Employees, skin conductance was a more effective way to share knowledge about customer experience when compared to traditional story telling.

By answering design questions with a variety of industries, I helped establish thick psychophysiology as a way to conduct explorative psychophysiology. Interviews, observations, and grounded theory, helped color in the context of psychophysiology and allowed me, as a design researcher, to better understand customer experience.

1.4 OVERVIEW

This dissertation reflects the split in my own interest: rigorous science and applied interventions.

Chapter 2 describes the specificity problem: a physiological response can mean many different things, Specificity challenges are inherent in ambulatory studies where stimuli and environment are uncontrolled. This need for specificity underlines the necessity and usefulness of thick psychophysiology in exploratory research. In addition to outlining the specificity problem, I review other possible methods for improving specificity in ambulatory settings.

Chapter 3 provides a detailed summary of how I conduct thick psychophysiology. This chapter outlines a variety of challenges with interpreting ambulatory data.

Chapters 4 and 5 show results from two studies done with thick psychophysiology. I describe the findings and their importance from two case study projects working with the New World Symphony and the LEGO Group®. I write these case studies in an accessible way that everyday employees can understand. I then describe how the data was analyzed and findings were inferred in an additional section.

Chapter 6 describes a survey of 143 LEGO employees who were asked about a new initiative. We demonstrate how video with skin conductance can help increase employee's support for this new initiative better than traditional design stories or video. Further analysis explores the interactions between people who respond empathically and those who viewed skin conductance. We also discuss the impact of describing skin conductance in an open-ended manner compared to a simplified explanation.

Chapter 2: The Specificity Problem

In the last decade, with the onset of smaller, more sensitive, and new types of technology, ambulatory measurements have become a feasible tool to measure changes in psychophysiology (Fahrenberg, 2006) along with other responses or features of a person in daily life. A person's cardiovascular activity, respiration, EMG, or skin conductance can be measured outside the laboratory, 24-hours a day. These measurements can help determine whether lab findings extend to people's everyday lives and are externally valid. In addition, they can help reveal some phenomena that may be difficult if not impossible to observe in a laboratory. While ambulatory methods can provide new ways of understanding physiological responses, this technology also brings new challenges: now the "experimental settings" are no longer fully-planned or well-controlled. Ambulatory methods can have low specificity, meaning that we may be unable to determine what factor(s) resulted in a specific physiological response. In the second part of this paper, we discuss what we will call the "specificity problem" in ambulatory physiological measures, and then consider how future ambulatory studies may be able to address the specificity problem. Ambulatory methods more broadly include psychophysiological measures, daily diary methods (also called electronic momentary assessments or experience sampling), and measurement of other aspects of the person or environment (e.g., posture or ambient temperature); however, we will focus here on ambulatory physiological methods used to address research questions about the psychological state of the participant. We will also focus on accompanying measures of the person or environment that are needed to enhance the ability to make inferences about the possible relationships between psychological state and physiological change[2]. Most ambulatory examples will be from heart rate or skin conductance due to their dominance in the ambulatory literature; however, the concepts presented can apply to other ambulatory measurements as well.

2.1 AMBULATORY MEASURES PROVIDE EXTERNAL VALIDITY AND ALLOW FOR BROADER OBSERVATIONS

Some physiological measurements obtained in the lab have shown little or no resemblance to real-world situations (low external validity). Wilhem and colleagues

[2] Holter monitoring and ambulatory BP are also used in health studies as well, which will not be addressed.

summarize the benefits of ambulatory settings (2010). Laboratory settings are not exact replicas of the real world. Individuals arrive to a lab and are given stimuli that may have little or no meaning to them. The lab is contrived; subjects must sit still and are placed in an unfamiliar context. Participants may recognize the situation as not real, compromising their reactions.

Whether or not lab results generalize to real world settings is typically unknown. For example, patients' measured-in-clinic blood pressure and heart rate can differ from typical every-day cardiovascular output. Some patients show "white-coat hypertension," in which their blood pressure is elevated, likely due to the anxiety of a medical exam (Kamarck et al., 2003). Other patients' blood pressures go down during a medical exam, known as masked hypertension. While a patient may actually have hypertension in day-to-day life, the medical setting can prevent doctors from observing that patient's everyday blood pressure. In these cases, the clinic may have caused less strain than a daily living environment (Pickering et al., 2007). In both examples, monitoring ambulatory blood pressure across an entire day provides a better indicator of hypertension (Mallion et al., 1999). As another example, laboratory research suggests a relationship between panic disorders and a reduced respiratory rate (Wilhelm et al., 2001); however, Pfaltz and colleagues (2009) found respiratory rate did not statistically differ between people with panic disorders and controls in real world settings. Differing results between laboratory and ambulatory research have also been shown with cardiovascular responses (Kamarck et al., 2003; van Doornen et al., 2009).

Many day-to-day phenomena are difficult, if not impossible, to observe in a laboratory setting. Important life events, stress at work, a fight with a spouse, time spent with friends, etc., are challenging to measure in the sterile environment of a fully controlled laboratory, especially with fixed stimuli. Looking at a picture of a rollercoaster is a different sensation from riding a rollercoaster. In an attempt to compare lab and real-world physiological stress responses, (Wilhelm & Grossman, 2010) measured a woman's heart rate across an entire day, controlling for physical activity. During the morning, the subject was given five mental stressors in a laboratory setting: reading, simulated public speaking, auditory attention task, memory comparison reaction time task, and a math task. A small (5-10 bpm) increase in heart rate was observed during these stress trials. Later, in that same day, the subject watched her national soccer team on TV as she sat down. During this natural event, her heart rate increased by over 50 bpm. The emotional, ambulatory event dramatically increased heart rate, whereas a battery of laboratory tests were unable to produce similar results. By measuring physiological signals across the day, we can begin to observe how subconscious, psychological processes occur in daily lives, opening up a richer set of data for exploration.

2.2 SPECIFICITY WITH AMBULATORY SETTINGS

Ambulatory measurements provide a needed toolkit for psychophysiology. However, these measurements have low specificity, making interpretation difficult. Specificity is the number of psychological or physical changes that map onto a physiological response. Typically, a physiological response can occur for a variety of reasons (many to one). In order to make strong inferences, scientists must be able to map exactly one psychological response to exactly one physiological response (one to one) (Cacioppo & Tassinary, 1990). For example, many events can increase electrodermal activity, but for an ideal interpretation, skin conductance responses will increase with only one psychological response (Figure 2.1).

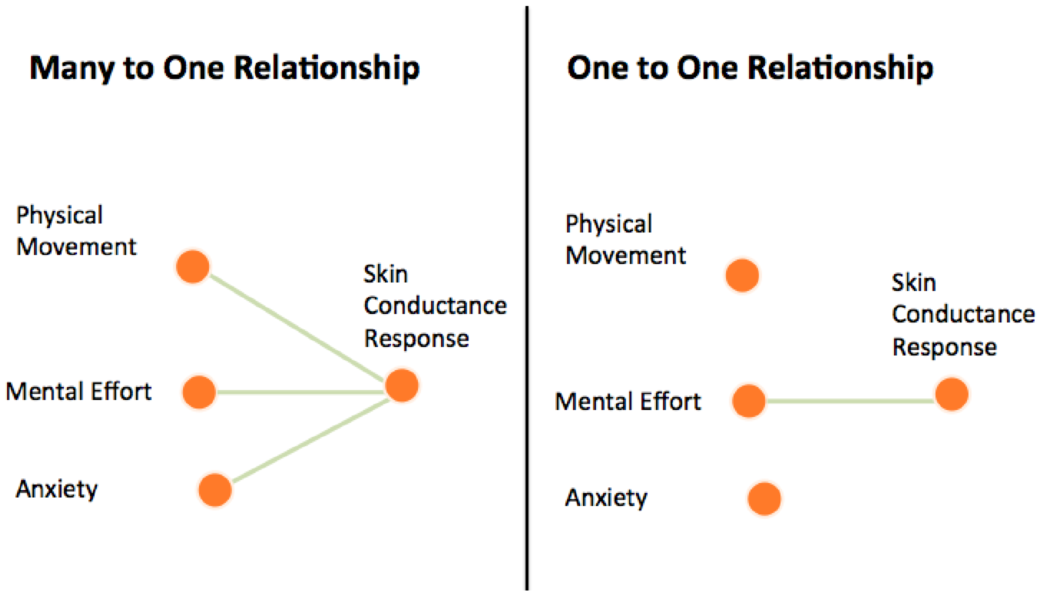


Figure 2.1 Skin conductance responses occur for a variety of reasons. For inference, a one to one relationship is needed. Adapted from Cacioppo and Tassinary (1990).

If researchers fail to create a one to one relationship (high specificity), then conclusions are limited. In our simplified example, a skin conductance response may occur because of physical movement, mental effort, or an anxiety response. This lack of specificity makes linking the skin conductance response to a specific psychological trait difficult. Both laboratory and ambulatory studies need to account for specificity, but ambulatory research can often have a larger range of possibilities for why a response may have occurred.

As an ambulatory example, Springer and colleagues (1989) attempted to measure the mental workload of 33 students while working with computer aided design software (CAD) and an elevated drawing board. Skin resistance response amplitude was larger while the students stood at the board. However, increased workload may not be the cause of the amplitude change. The authors noted how skin resistance could change with physical movement at the drawing board. As skin resistance could increase from either mental workload or physical movement (many to one), whether CAD helps reduce mental workload cannot be determined from skin conductance alone. The researchers emphasized other research measures because of this specificity problem.

As another, more complex example, data of cortisol fluctuations in day-to-day settings can be difficult to interpret. Laboratory results have suggested that cortisol is elevated when an individual is stressed (Pollard, 1995). However, a review of field studies demonstrated cortisol might have a positive, negative, or no correlation effect with self-reported stress levels at work (Hjortskov et al., 2004). The authors suggest this difference may be in part due to uncontrolled factors such as depression, negative mood, infectious diseases, alcohol intake, diurnal variation, or use of contraceptives. Adam and Gunnar (2001) observed a correlation between hours worked and cortisol levels for mothers. The authors questioned whether the correlation between cortisol levels and work hours was due to the work environment or the stress interacting with their children with limited hours at home. Rather than one controlled factor, a whole host of factors influence cortisol levels in the field, reducing the specificity of cortisol and limiting possible conclusions.

Laboratory experiments also have specificity challenges (Cacioppo & Tassinary, 1990), albeit more easily controllable. Researchers follow a multitude of rules when conducting psychophysiological experiments. For electrodermal activity, subjects cannot move their body, posture, hands, or fingers, nor do any other physical work[3]. Participants are asked to breathe regularly without talking. Any external stimuli needs to be removed from the setting: no noises, no talking, no food, and no surprise appearances. The environment needs to be standardized as well: no changes in lighting, temperature, humidity, or other environmental factors. Any additional stimuli or environmental changes could result in a change in skin conductance response, making the cause of the skin conductance response unknown. Controlling for these responses inside a lab is relatively easy: have all subjects come to the same laboratory where external stimuli are blocked out. Instructing participants

[3] Though some studies allow for a small amount of manual activity such as pushing a button or using a mouse.

not to move, to talk, or to breathe hard eliminates most behavioral factors. Experiments can be videotaped in case participants do move or talk, so the faulty data can be removed.

Laboratories control for psychological changes as well.[4] Electrodermal activity can increase from anxiety, mental effort, positive mood, or negative mood. To allow for inference with electrodermal activity, researchers carefully choose the stimuli to be presented. If a participant is told to prepare a public talk, the assumed stimulus is anxiety, with participants being asked about their anxiety afterwards (Kirschbaum et al., 1993). Lang and colleagues tested the difference between positive and negative images; they showed carefully selected photos that viewers previously rated as positive or negative. Immediately after viewing the images, subjects self-reported their affective state (Lang et al., 1993). Thus, to verify the elicited mental state, researchers rely on self-reports and normatively rated stimuli.

Many times, ambulatory measurements do not have the benefit of controlled settings and lack control over some factors that may have an influence: participants will move around, vary their environments, and be exposed to a multitude of stimuli that are not well-defined. These uncontrolled settings can leave a researcher in the dark for why a change may have occurred. In ambulatory measurements self-report data is often absent and stimuli are novel and unique, making inference difficult.

Although ambulatory research can have low specificity, there are ways in which researchers can either create the needed one to one relationship or infer results without specificity. In the next sections we suggest a variety of ways in which ambulatory measurements can increase the likelihood of making a strong psychophysiological inference, i.e., reduce the "specificity problem".

2.3 LISTENING TO A CLASSICAL CONCERT AS A CASE STUDY

As an example, we focus on one project we have been working on: measuring the attention and engagement of participants as they attend a classical concert. We measured the electrodermal activity of 40 participants from their middle phalanx as they viewed a classical concert at the New World Symphony. Afterwards, we conducted 1 to 2 hour-long interviews with subject pairs. During these concerts a variety of environmental challenges were

[4] Not all thoughts and emotions can be controlled for in laboratory experiments. Participants can become bored or anxious waiting for an experiment, participants may be thinking about their exam in an hour rather than focusing on the screen, and individuals can interpret stimuli differently. So while laboratory experiments provide more constrained stimuli, psychological responses can still vary.

presented – ringing phones, clapping hands, talking, posture changes, and more. However, we were able to show a strong emotional response in anticipation of and listening to *Romeo and Juliette* (Hedman et al. 2013). Below is Anita's skin conductance response to this concert.

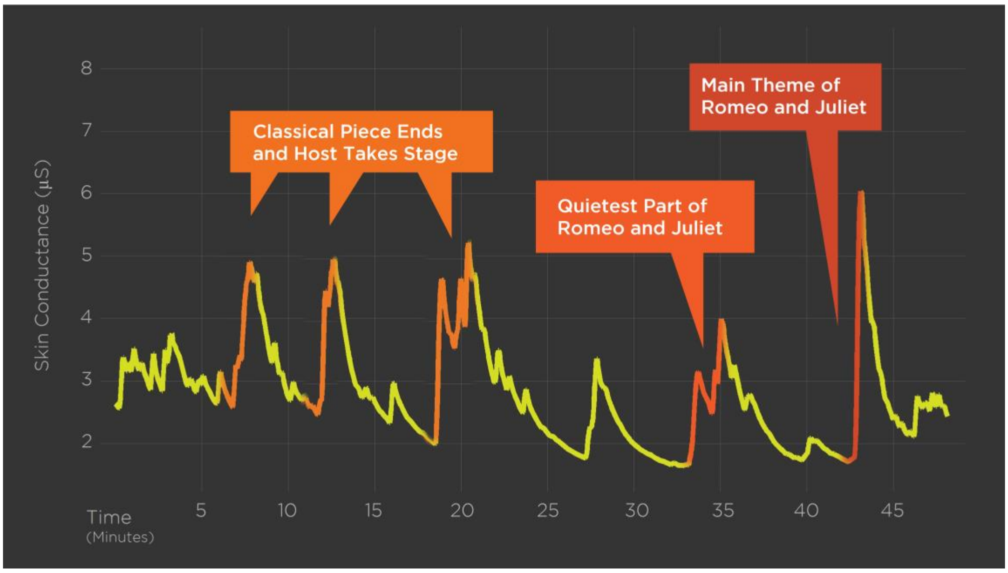


Figure 2.2 Anita's skin conductance changes across a classical concert, but how can researchers interpret these responses?

While the majority of analysis will focus on Anita's complex response, we also measured three other people's responses during these sessions as well. Below all four people's responses are graphed simultaneously.

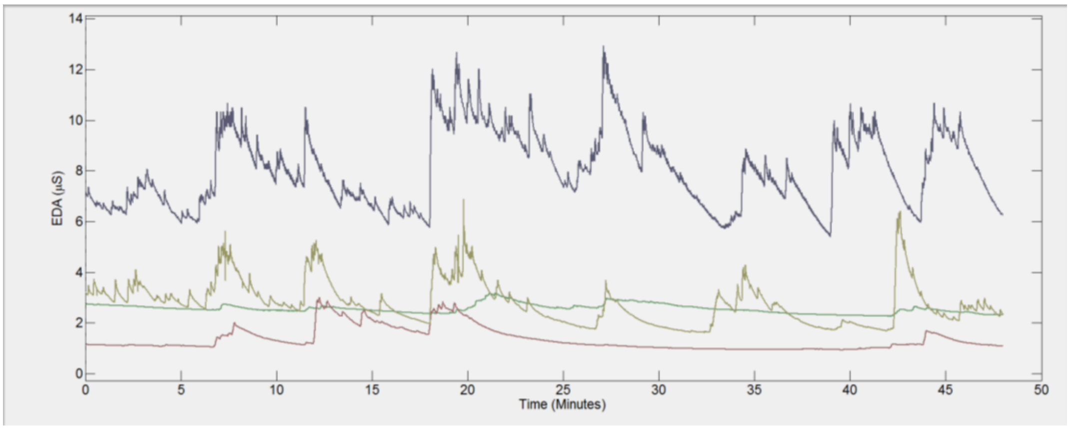


Figure 2.3 Four people's skin conductance responses to the Romeo and Juliette Concert. Anita's response is the yellow one.

Some of Anita's skin conductance responses occur at times when other participants do not respond. At other times, all four participants respond near the same time, like the large response to a transition around minute 18.

2.4 USE OF MULTIPLE MEASUREMENTS

The primary way of creating specificity in both controlled and ambulatory research has been incorporating multiple measurements. Rather than keeping factors constant, researchers can measure and account for additional factors. We detail a list of common factors and discuss ways of accounting for their influence in ambulatory settings.

2.4.1 Physical Influences

Micro-Movements

When participants touch or accidently adjust electrodes or the measuring device, data can be altered in large ways. For example, an extra response from a touched electrode can have large effects on calculated heart rate and heart rate variability (Stuiver & Mulder, 2009). When participants put pressure on the electrodes or the tape becomes loose, we will see large responses and drops in signals that could appear as an additional skin conductance response. Many electrodes do not take much effort to manipulate. For electrodes on fingers, an individual moving their knuckles, flipping through a program, or resting their hand on their knee could all produce artifacts.

One way of detecting these changes is to detect their presence in the signal. For example, when an electrode becomes loose, a rapid decrease occurs, which does not happen naturally (see 3.10.3 for example of artifacts). Alternatively, researchers could use force sensors (Lee & Nicholls, 1999) to measure the current pressure applied to the electrodes. If participants are videotaped, movement where the electrode is placed could be a marker for artifacts as well. We measured skin conductance on the left and right side of the body simultaneously and removed data that occurred on only one side, as these changes may be due to micro movements (Hedman, 2010). However, this solution may be ill-suited if differences could naturally occur between locations. For example, skin conductance has been found to be under ipsilateral control by several subcortical regions (Mangina & Beuzeron-Mangina, 1996). Picard et al. (In Press) observed asymmetrical differences in skin conductance in both controlled experiments and day-to-day activities, so some psychological effects could be missed by only considering moments where the two sides concur. In the future, new ways may be developed to secure sensors (Debener et al., 2012), but these methods should be tested as even a taped down electrode can slip with enough participant movement.

Metabolism and Movement

Changes in metabolism can have a large effect on physiological signals, much larger than corresponding psychological states (Blix et al., 1974). When an individual lifts a hand or a foot, there can be large skin conductance responses. Exercise like moving up a hill or stairs can produce physiological changes as well. In our symphony work, transitions between songs corresponded with large responses in skin conductance. These musical transitions also included clapping that can create skin conductance response as well. In our initial trials, we had to ignore transitions between songs because of this movement. In later sessions we asked participants to refrain from clapping and still saw large responses during transitions – suggesting there is a psychological response as well.

One way of measuring changes in metabolism is by measuring oxygen consumption either directly (Blix et al., 1974) or in a calibrated band across the torso and abdomen (Wilhelm & Roth, 1996). As sensors have become smaller, movement and corresponding metabolism have been measured with accelerometers (Myrtek, 2004; Poh et al., 2010). During treadmill activities, accelerometers strongly correlate with oxygen uptake (the Tritrac-R3D had a correlation of 0.93); however, these correlations are substantially decreased in other lifestyle activities. For example, the Tritrac-R3D had a correlation of 0.59 during lifestyle activities (Trost et al., 2005).

Movement may change a physiological signal without changing metabolism as well. When individuals scratch themselves, or move their arm from an armrest to a knee we have observed changes in skin conductance. These movements are not likely to have a large effect on metabolism but can sometimes produce large skin conductance responses. Researchers could put accelerometers on study participants' hands, but would then need to put corresponding sensors on feet, heads, and other places where twitches could occur. In our work, we rely on video tape to remove these physically induced responses. During dynamic interactions – someone talking with their hands, trying out retail products, writing a paper – these movements can become too numerous to account for. Additionally, one can conduct hard work or muscle movement without much visible movement – imagine an individual clenching a fist or carrying a heavy backpack.

Posture

Like movement influences, changes in posture can also have a large effect on physiological signals as well (Pomeranz et al., 1985). When participants at the symphony would shift in their chairs, bend over to pick up a program, or stand up, skin conductance responses could occur. Strain gauges could be worn to signal changes in body position (Lorussi et al., 2004). Additionally, a network of accelerometers can help determine body position as well, taking advantage of gravity's natural effect on accelerometers (Foerster & Fahrenberg, 2000). However, determining position based on accelerometer data becomes more difficult as an individual begins to actively move (Hansson et al., 2001).

Breathing, Yawning, Sneezing, and Talking

Taking deep breaths can also change a physiological signal (Hirsch & Bishop, 1981). One of the ways to produce a reliable skin conductance response is to have participants breath hard or blow up a balloon (Lykken & Venables, 1971). Effects like coughing, yawning (Greco & Baenninger, 1991), and sneezing (James & Daly, 1969) can affect physiological responses as well. A simple way of detecting changes in breathing is with a chest strap and removing data that co-occurs with abnormal breathing patterns (Bearden & Moffatt, 2001).

The process of talking can be an emotional experience and is often a subject of study; speech has been shown to affect psychophysiology (Spitzer et al., 1992; Westenberg et al., 2009). Talk in itself alters breathing, which in turn can affect physiological signals, making interpretation more difficult. One solution for measuring psychophysiology during speech is to only compare times when subjects are talking at similar volumes and rates.

Summary for Movement

Physical changes in the body can have a large effect on a person's physiological state. Often times, measuring and taking into account exact changes in physical activity is impractical. In these cases, researchers should be vigilant in interpretation as physical changes can substantially alter the data and alter interpretation. For our symphony work, we saw large increases in skin conductance during intermissions and clapping, however we removed those spaces from analysis as there was a change in physical movement. However, when the subjects were sitting rather still during the concert, we could conclude that physical movements and changes were not likely the reason for responses. Participants could still have moved their feet or began rapidly breathing, so we relied on results from multiple subjects to account for these random influences.

2.4.2 External Influences

A laboratory setting is a relatively constant environment with the main external stimuli being controlled by the researchers. However, in the real world, the environment is constantly fluctuating around us and we are constantly responding to these changes. In this section, we describe ways of accounting for environmental fluctuations.

Temperature/Humidity

Change of location or time can introduce changes in temperature and humidity. These changes can affect physiological responses, in particular skin conductance (Waters et al., 1979). Testing responses both in and outdoors or during morning and night could be affected by temperature and humidity. Sensors today can easily measure temperature and humidity (Fletcher et al., 2010; Meng et al., 2004; Meng et al., 2004).

Time of Day

Physiological responses can change diurnally (Hot et al., 1999; Yamasaki et al., 1996). Most sensors today have a clock function, and studies should account for time of day. Ideally, timing is recorded in relation to circadian cycle as well, since many physiological processes are altered by sleep regularity.

Environmental Stimuli

Physiological responses are sensitive and can respond to many different stimuli. While these responses might be psychologically relevant, the causal mechanism may not be observed, thus leading to misinterpretation. For example, during some of our symphony views, a large group of people came in late, clacking their heels. Afterwards, participants reported feeling annoyance toward these guests. The skin responses during this heel clacking could occur for two different reasons. First, a new sound was heard, likely creating an orienting response to novelty (Lynn, 1966). Second, the users reported frustration, which can also change skin conductance (Scheirer et al., 2002). If we were to miss this small event, the large skin conductance response could easily be misinterpreted as a response to the music or concert itself. These environmental changes are occurring constantly. Some of the stimuli we noted at the concert included a touch on the back by a husband, scratching oneself, eating a snack, drinking alcohol or coffee, chewing gum, a triangle player messing up and catching the listener’s attention, reading a program, falling asleep, and noticing a nearby person sleeping. This long list is to emphasize that even during a one hour concert, where stimuli are discouraged, the number of stimuli present was still large. Wilhelm and Grossman outline a series of these contextual variables and suggest methods that may take some of them into account (2010).

To "sense" all contextual influences would be a substantial undertaking. One could imagine a detector on the mouth to detect drink items, touch sensors covering the body in case of bodily touch, a microphone to detect any audio conversation, etc., but these measures are impractical today. For the symphony, we video and audio taped the concert and also observed the participants in the setting, which helped us catch some of these stimuli, but not all of them.

2.4.3 Psychological Influences

While most ambulatory physiological measurements aim to be a marker of a psychological state, determining what psychological state was produced can be difficult as physiological responses can occur from a variety of mental states.

Drifting Mind/Day Dreaming

During the concert, some people reported that their minds drifted off. They began thinking about what was for dinner or what to do after the concert. These thoughts in themselves can be exciting, perhaps even more exciting than the concert, producing a skin conductance response. It may be the case that mind wandering is quite pervasive, occupying 15-50% of participant's time (Smallwood & Schooler, 2006). By looking at ocular motility, it may be possible to detect when an individual is daydreaming (Singer et al., 1971). Individuals could also report when they have a "task-irrelevant thought" by pushing a button (Antrobus et al., 1967), though the accuracy of correctly reporting daydreaming should be studied more.

Mental Work

When a person is presented with a difficult mental task, physiological signals can also change (Kramer, 1991). A person may not be experiencing an emotional event but rather thinking hard. During the concert, were the new participants excited about the concert or conversely mentally working hard to listen and understand the complex music? Mental workload could be detected externally with a survey or behavioural measurements. For example, we noted when the skin conductance of a child with Autism Spectrum Disorder rapidly increased as she planned how to climb on top of a swing (Hedman et al., 2012). We knew the process of motor planning for this child would be difficult, as the therapist had informed us. For the symphony example, we relied on self-reports, noting times when participants reported being confused or thinking hard.

Affective State

If all other influences are correctly accounted for, identifying the affective state that triggered the response can still be challenging. Skin conductance can increase for both positive and negative valence stimuli and for other mental states as well including disgust, fear, happiness, and sadness (Lang et al., 1993). When a subject responds to the *Romeo and Juliette* is she excited or stressed? We use self-report to evaluate these responses. When asked about the theme, Anita waved her arms to her chest and told us how wonderful it was. It may be possible to detect discrete emotional states with additional physiological responses also, which we cover in the Measuring Unique Signals section.

2.4.4 Additional Challenges to Using Multiple Methods

Correctly Accounting for Influences

Once an influential factor is observed, the physiological response needs to be correctly adjusted. Physiological responses often are non-linear, and correctly subtracting or adding values to a signal is complicated and understudied. For example, increased metabolism corresponds with an increased heart rate.

One way of accounting for physical movement's affect is additional heart rate (Blix et al., 1974). Muscular workload during exercise is proportional to O_2 consumption, so O_2 consumption can predict what an individual's typical heart rate would be. The emotional component of heart rate (the additional part) can be calculated by subtracting this predicted heart rate (based on O_2 consumption) from the current heart rate. This method requires a previous test (usually more than one) to document the relationship between the physiological signal and influential factor for each individual and assumes this relationship will remain constant across time.

A more conservative method is to only compare data when factors are constant; for example, we can compare physiological measurements at times when other influential factors are the same. By comparing times when physical activity is similar, one can assume physical activity was not a major influence relative to other situations with the same extent of physical activity.

A third option is to remove data from analysis when an additional factor is present. When subjects moved their bodies or clapped during the symphony concert, we ignored that data, as we were unsure how electrodermal activity changes with these factors. Researchers also need to set the right measurement epoch. In Anita's data, it takes four minutes for the skin conductance level to return to pre-clapping levels. Removing data makes interpretation more difficult; the amount of time an individual is not moving their hands or body, not talking, and not influenced by an external stimuli cuts out a large portion of data in our symphony work.

Interaction of Factors

Additional heart rate can help when movement is the only other variable of interest; however, when more than one factor influences an individual at the same time, interpretation becomes more complicated. In the additional heart rate example, the baseline relationship must remain constant across time, which may not be the case. For example, if an individual drinks a cup of coffee with caffeine, this may change their heart rate afterwards (Green & Jordan, 2002). The relationship between oxygen intake and heart rate may change because of the caffeine, and this change may not be linear. The interaction effects between variables are complex and laborious to model especially for more than three factors.

Overall, this section underlines the potential and also large number of challenges to correctly using multiple measurements to narrow a physiological response for a one to one mapping. This section was intended to help orient ambulatory researchers to the challenges they may find in interpreting results and potential areas where more research will be helpful. Below we outline other ways of either increasing specificity or being able to infer without specificity.

2.5 COMPARE GROUPS OR USE A CONTROL GROUP

Laboratory research controls for the large number of factors that create signal variability by comparing across groups. Additional factors can often be thought of as noise that influences both groups equally. For example, Adam and Gunnar (2001) compared cortisol levels between mothers who worked long hours vs. a shorter number of hours. Some factors, such as food eaten, amount of exercise, and time outside likely did not significantly differ across groups and therefore were probably not a cause for observed cortisol differences. However, the authors still needed to account for any additional variables that could be influenced by work hours, like wake-up time and relationship status. The list of influential factors that may be the real cause for differences (third variables) with ambulatory measurements – different groups may be engaging in differing physical activities, exposed to different environmental influences, and experiencing different psychological states.

2.6 MEASURE UNIQUE RESPONSES

Finding a one to one mapping of a skin conductance response to a psychological response is most likely impossible. However, rather than mapping basic physiological responses, researchers can look at more complicated relationships between different signals and time. Cacioppo and Tassinary give the simplified example of differentiating between orienting, startle, and defensive responses from individuals (1990). All three responses produce a skin conductance; however, heart rate only increases during startle and defensive responses. The heart rate responses between defensive and startle responses also differ, with heart rate rising and peaking much later with defensive responses (Lynn, 1966; Turpin, 1986). By analyzing skin conductance and heart rate responses across time, researchers can differentiate the three responses. By combining signals together and looking at their patterns over time, we could find unique responses that only occur during specific psychological responses.

Some researchers are using modern pattern recognition procedures to classify psychological states from multiple physiological signals (Christie & Friedman, 2004; Kolodyazhniy et al., 2011; Kreibig et al., 2007; Stephens et al., 2010). However, these multivariate methods are often done with only one controlled stimulus like a predetermined movie or music, and future work will be needed to determine if these effects are consistent across environments, which would be essential for ambulatory studies. (Quigley & Barrett, 2014). In one ambulatory, neurological study, Poh et al. identified times when children were having a seizure based off of a combination of accelerometer data from the wrist with skin conductance data (2012). By analyzing complex interactions across signals, it may be possible to identify unique responses that have a one to one mapping with a psychological state.

2.6.1 Filter Signals

Similar to finding unique responses, researchers can selectively pay attention to signals that pass certain thresholds. While skin conductance responses occur throughout the day, large skin conductance, with amplitudes greater than 2 uS occur fewer times. In analyzing the symphony data, we filtered responses to only those that were in the top 10% of increases in amplitude (Hedman et al., 2012). Small changes that may have occurred because of small movement shifts, novel stimuli, or for a non-specific reason were ignored. We were left only with points where a more influential event occurred. These filters do not create a one to one mapping on their own; however, the largest skin conductance responses we observed in the symphony study occurred during movement at the intermission and ending, which would still need to be accounted for. Filtering out signals will require more research on what range of values psychological changes can alter a physiological signal, and what range of values other factors can also alter that signal. It may be the case that the other factors have a much larger range than the psychological state, making filtering difficult.

2.7 MEASURE ABSENCE OF RESPONSE

Physiological responses can occur for a variety of reasons (many to one). However, the inverse of this relationship is true as well, if a physiological response does not occur, than all the causal mechanisms that produce that response likely did not occur as well (assuming no measurement issues are present and the subject is not stabile). An absence of response can be interpreted as an absence of stimulus, bypassing the "specificity problem" (Cacioppo & Tassinary, 1990). Researchers in marketing have noted times during ads when participants do not physiologically react to an ad, suggesting those ads may not be as activating (Hazlett & Hazlett, 1999). In the symphony example, during a song performance from minute 24 to 27, no skin conductance responses occurred for Anita. Since the process of orienting can produce skin conductance responses, we conclude that she did not likely orient or emotionally respond to the plucking string, long pause, or crescendo in this section.

2.8 USE CONDITIONAL PROBABILITY TO IDENTIFY RELATIONSHIPS

Rather than treat physiological signals as a marker of a psychological state, researchers can show relationships between physiological signals and psychological states. Using a conditional probability paradigm, one can determine a relationship between variables by analyzing their covariance. For example, Myrtek (2004) tested how increases in additional heart rate corresponded to various emotional states. Whenever additional heart rate increased past a threshold, the participant was signaled to select their current emotional state. Participants were also polled at random time intervals. For analysis, Myrtek compared the frequency of reported emotions during triggered and randomized times. The probability of a person being in an emotional state when he has increased heart rate is compared to the probability of that participant reporting the emotional state during randomized times. A difference in probabilities would suggest that increased heart rate correlates with the measured emotional state. As another example, Cacioppo and colleagues (1988) measured students' EMG over the brow region while they spoke. Specific patterns of EMG corresponded closely with the individual's self-reported emotional state. With a conditional probability paradigm, researchers can make better inferences about the relationships between physiological changes and the accompanying psychological state.

2.9 DO NOT ASSUME SPECIFICITY

In every ambulatory study we presented, physiological changes could occur for a variety of reasons, many unaccounted for. While laboratory experiments often design their experiment to allow for specificity, ambulatory researchers should often assume low specificity; the physiological responses could occur for a variety of reasons. Often times, bringing a physiological response to a one to one mapping is too limiting or impossible. However, ambulatory research can still be useful when the lack of specificity is acknowledged.

Fahrenberg (1996) suggested that ambulatory measurements were a compliment to laboratory experiments. Ambulatory research can confirm the external validity of laboratory results and also help researchers discover new psychophysiological relationships. These new discoveries can later be tested in laboratory settings where specificity is higher. In the symphony example, we noticed that Anita's skin conductance changed during transitions between pieces. Novel stimuli increasing skin conductance has been confirmed in laboratory settings multiple times – novel stimuli create responses and this study showed this effect in an ambulatory setting. We also noticed Anita showed a large skin response to *Romeo and Juliette*, a song with which she was familiar. A laboratory experiment could be designed to further test whether familiar classical music creates larger responses than unknown music.

2.10 CONCLUSIONS

In the future, ambulatory psychophysiology measurements will become more and more accessible - allowing scientists to understand psychological responses in real world settings. These tests have the potential to help show the external validity of theories and provide a ground for new types of questions that cannot be answered in controlled settings. However, ambulatory measurements must acknowledge the "specificity problem" or inferences may be incorrect.

Using multiple measurements is the go-to approach to account for influential factors, but measuring all of the possible additional factors can prove challenging and in some settings nearly impossible. Like accounting for movement with additional heart rate, researchers can begin to measure and account for other changes like breathing and changes in posture. An individual’s psychological state and external influences like talking to a person or a cell phone ringing are harder to measure, but could be accounted for with new measures. Future research should continue to develop new ways of measuring these additional factors that influence psychophysiology and develop new ways to best account for these factors in analysis.

We also discussed other ways researchers can make conclusions without having a one to one mapping. A conditional probability paradigm can help establish relationships between psychological states and psychophysiological responses. Myrtek (2004) compared probabilities of emotional states during increased heart rate and randomized times. Cacioppo et al. (1988) showed correlations between emotional states and facial EMG responses. Future researchers can compare the probability of responses during changes in physiology or controlled times as a way of testing a relationship. This method has the advantage of being agnostic to other influential factors, as probabilities are compared.

Identifying moments of lack of response also can provide psychophysiologists with inferences, even with low specificity. While most ambulatory research measure when responses occur, research about when responses do not occur may provide valuable learnings for questions where specificity is low. When participants in the symphony response did not produce skin conductance responses, we were able to conclude that they were likely not orienting or emotionally responding to the music, despite multiple variables capable of increasing their skin conductance in that setting.

The final option is that researchers do not assume specificity. While inferences are reduced without specificity, ambulatory measurements can still provide new information that can be further tested in controlled settings. For many of our questions, the symphony research had low specificity, yet the research also revealed interesting physiological responses that could be tested and refined further in more controlled settings.

This chapter is a beginning road map, a list of strategies we have personally used in interpreting the messy world of ambulatory methods. As the field grows, we look forward to new methods and techniques to either work around the specificity problem or address it. There is much to find in the real world from ambulatory methods, allowing researchers to test real-world reactions in meaningful contexts. With the right critical eye and methodological forethought, ambulatory research can provide new insights and inferences helping to bring the every-day world to psychophysiology

Chapter 3: The Thick Psychophysiology Method

3.1 BACKGROUND

In this chapter, I describe how to make inferences from exploratory research with a method I entitle "Thick Psychophysiology." Thick psychophysiology uses observational methods and interviews to gain a thick description of an event occurring in the real world. I define thick psychophysiology as having these four characteristics:

1. Quantitatively measures a physiological aspect of an individual.
2. Answers an exploratory, ambulatory question.
3. Researches external context: what events lead to a physiological response occurring?
4. Researches internal context: what is the meaning and interpretation of the physiological response?

In Kara's therapy example, the external context includes information like: the therapy session was early in the morning, the child mentioned "wagon" multiple times, and the child had to listen to her own voice. The internal context describes the child's emotional reaction – the therapist reported that the child appeared overwhelmed or the child tried to take the headphones off, suggesting she may be frustrated.

While thick psychophysiology itself does not have a background of practice, various fields of research have implemented a combination of the described elements, and are worth noting to help describe thick psychophysiology's contribution and inspiration.

3.1.1 Traditional Ethnographic Methods

Anthropological and sociological research is rich with observations and interviews about the nature of emotion. Rosaldo lived in the Philippines for two years and noted how the Llongot tribe's feeling of shame differed from Western shame, and how that shame played a large role in young males while headhunting (1983). Hochschild attended a Delta Flight Stewardess training and interviewed numerous expert and training flight stewardesses to understand how a flight stewardess's emotional responses are commoditized – "service with a smile" (2003). Katz videotaped French tourists for 8 hours as they interacted in a French, mirror fun house (2001). He analyzed frame by frame when people laughed and how others around them behaved. He noticed that visitors hardly ever laughed alone, and when they did, it was often to gain attention of colleagues.

A review of all the observational studies about emotions done from an anthropological or sociological perspective would be a feat in itself (e.g. Leavitt, 1996; C. Lutz & White, 1986; C. A. Lutz & Abu-Lughod, 1990), but one consistent pattern is that the research does not use psychophysiology. The few times sociologists and anthropologists have used physiological measurements, the research resembles laboratory experiments rather than exploratory research. For example, Cohen and colleagues (1996) studied how aggression occurs in southern United States' culture compared to students who grew up in the northern part of the country. A confederate in the hallway bumped into the participant and called the participant an asshole. Quickly after the pre-planned event, participant's testosterone and cortisol levels were taken and differences between physiological reactions were determined. This study did not develop thick descriptions and tested a set hypothesis. Rather than explore how southern culture affects emotional reactions with observations and interviews, these researchers tested a previously held theory about southern culture, much like traditional psychophysiological laboratory experiments.

3.1.2 Laboratory Settings in Ambulatory Setting

In order to establish external validity or measure more "real world" responses, the majority of ambulatory psychophysiological research tests specific hypotheses (for a collection of examples see Fahrenberg & Myrtek, 2001, Fahrenberg et al., 2007). Researchers have different groups of people participate in predetermined situations and note the differences in responses. Boucsein and colleagues measured NS-SCR reactions for 24 hours (2001). Half of the participants heard a neutral story while the other half heard a horror story right before they went to bed. After subjects heard horror stories, their NS-SCR frequency and NS-SCR amplitude increased. Wilhelm and Roth had two groups of participants, flight phobics and controls, all fly simultaneously on a plane (1998). He noticed that flight phobics had elevated skin conductance from the left hand palmer surface and increased additional heart rate. In this type of research, the main focus is to validate the tools as usable in ambulatory settings. The researchers even named their paper "Taking the Laboratory to the Skies." This data could have been explored further. For example, the responses of the graph show that phobics and controls both increased skin conductance during takeoff, but only controls appeared to increase their skin conductance during the landing. However, these type of exploratory questions are seldom discussed in ambulatory psychophysiological studies.

3.1.3 Establishing Internal Context for Psychophysiology

When participants flew on the plane, Wilhelm and Roth asked participants three times how anxious and excited they were and what their desire to leave was. For phobics, self-report responses followed a similar response of heart rate and skin conductance across the flight with a peak during take-off. Without this additional measurement, it would be unclear if elevated skin conductance was connected to a participant's increased anxiety. Otherwise it may be due to their excitement, physical movement from turbulence, or other factors.

Psychophysiologists use interview questions as a way of gauging the internal context associated with an ambulatory physiological response. Thornton measured people's self-reported fear with a continuous response measurement dial and then showed the relationship between skin conductance and reported fear (2005). Kanning asked participants how they felt when movement was recorded (2012). Wilson measured the heart rate of pilots while flying and had them self-report their cognitive load as well (2001). Myrtek had participants report their emotional state throughout the day while additional heart rate was measured. Analysis showed how heart rate changed with different emotional states (2004). Cacioppo and colleagues interviewed participants with questions that varied in intimacy. Subjects were videotaped and their EMG was recorded (1988). Afterwards, participants described how they felt throughout the session. The researchers selected points with a specific type of EMG activity and had participants rate how they felt at those times on different scales.

Typically these surveys are quantitative with precise questions, which can be non-ideal for exploratory research. When developing thick description, ethnographic researchers tend to ask open-ended questions like "How did that make you feel?" or "What happened?" listening to how people describe their own state. With a constructivist's perspective, Daily asked children in a classroom to re-watch a video combined with skin conductance and describe what they believed the changes in skin conductance meant (2010). Mirza-babaei and colleagues noted times when skin conductance responses from left handed palmer surfaces occurred during video games, and then they interviewed participants about what happened at those times, noting times when game bugs occurred (2011). While Daily and Mirza-babaei both used open-ended questions, both studies were more about describing and validating a new method and did not share insights from participant's responses.

3.1.4 Establishing External Context for Psychophysiology

Even in controlled settings, psychophysiological responses are often influenced by external stimuli. From tones, to arousing photos, to preparing to give a talk, psychophysiological studies tend to use a contrived external event, which is expected to elicit a physiological response. For ambulatory responses, this external context can become more difficult to identify and categorize.

Taking a small step from controlled studies, some ambulatory research breaks down a physiological response into different time segments, and compares responses across these segments. Dozier and Kubak conducted attachment interviews with participants while measuring their skin conductance. People who reported being more hypersensitive also had higher skin conductance levels when talking about being separated, threatened, and having changes in relationships (1992). Wilson had pilots conduct a series of flight maneuvers and noted the largest increase in heart rate occurred when pilots were in the bombing range and during takeoff (2001). Rather than having a closed hypothesis, e.g., does flying increase skin conductance, the researchers sought out the times that a physiological response occurs, e.g. when does flying create the highest skin conductance responses?

Researchers can also note when specific events occur as a way of describing external events. Ravaja and colleagues noted every time players made a kill, were critically wounded, or died while playing a James Bond video game (2008). Those responses were correlated to skin conductance responses and facial EMG. Sands and Sands measured EEG of shoppers in a grocery store. Participants' responses differed when they looked at hedonic items like ice cream and alcohol compared to staplers and frozen food. Hazlett noted times during a racing game when participants passed another car or were passed, labeling these moments as positive and negative events and then noting how facial EMG differed between these events (2006). Healey and Picard (2005) had two researchers mark potential stressors of drivers by labelling when stressful objects like stop signs and bumps were present in a recorded video. By measuring when specific events occurred, researchers were able to better identify the external context around responses in uncontrolled settings.

When participants are all exposed to the same stimulus, like a video or an ad, changes in combined physiological responses can be connected to the consistent external stimuli.

Hazlett and Hazlett measured facial EMG to detect when participants smiled during ads (1999). Thornton measured electrodermal activity and self-reported fear as viewers watched ads about preventing car crashes. In one commercial, the largest skin conductance response occurred at the beginning when a pedestrian was hit by a vehicle. During a slow motion replay of the hit, the skin conductance response increased, but less dramatically. Daily and Picard noted which words participants were typing when physiological responses occurred in an electronic diary (2004). In these cases, researchers used changes in psychophysiology as a marker to highlight a consistent external event.

Researchers may also try to describe the situations around an event as well. For example, Kanning was studying how movement and affective state corresponded (2012). He asked participants to additionally report the context participants were in (working, leisure time, transport, and chores). From a thick description perspective, working only begins to describe an environmental context. Anderson and Lawler measured systolic heart rate reactivity as participants recalled a time they were angry (1995). Type A women showed greater reactivity when they expressed frustration of autonomy needs, and Type B women showed more reactivity when expressing frustration of affiliation needs. The authors specifically conclude the essential need to consider the context of the interview in order to find differences in responses.

I was only able to identify two research projects where explorative, open-ended observational methods were used to describe the external context. In the first study, Mirza-babaei had participants play a video game and recorded skin conductance from left handed palmer surfaces (2011). He then replayed parts of the video game where skin conductance response occurred and had participants describe what happened. For example, one player responded “I was not sure if I could still drive my buggy or if it was broken. I’ve started driving it again, but was not sure if it was going to explode soon or not. Eventually, it did.” Perhaps my favorite example of observational methods in psychophysiology was on accident. Zeier and colleagues wanted to test the heart rate responses of students when they participated in two tests and a lecture over three days (2001). However, the researchers were surprised when the largest increase in heart rate was not during these tests but at a time before hand. The researchers looked into what was occurring at that time, and discovered the students were practicing for an oral exam. He also noted on the third night, heart rate stayed more elevated than the two previous days. He determined this difference was due to the students going out to a party to celebrate the end of the three-day event. Zeier does not describe how he determined these events, but as we can see, the little knowledge of the external context he was able to produce helped him understand why a physiological difference may have occurred.

3.1.5 The Vision

When comparing Briggs's description of Eskimo's emotions to Zeier's description of a three-day academic event, psychophysiologists have a long way to go to produce strong, thick description. The work I present will be an attempt to bring exploratory research and strong observational methods to psychophysiology. I imagine a world in the future where psychophysiologists, anthropologists, and sociologists team up to understand day-to-day interactions with a grounded theory perspective.

This chapter is a detailed description of how a researcher might conduct thick psychophysiology. While the method seems simple, there are many complex challenges and obstacles that can occur in the process. I attempt to highlight these challenges and also show how my method has worked. This method has been reiterated and built upon over the past six years by measuring over 100 people's skin conductance and interviewing just as many. This method was used to help Lowe's and Best Buy improve their retail experience, the LEGO Group enter into the tablet space, IDEO solve various design problems, Hasbro learn about the stress of a new board game, and Google measure the emotional experience of the self-driving car. I cite specific examples where I can. In particular, as the retail experiences contain confidential specifics, I give the hypothetical example of a person buying a blender to help illustrate the methodology. I hope you find these suggestions informative and inspirational.

3.2 EXPERIENCE SETUP

When I started researching emotions, I treated people's responses like a typical circuit signal in an oscilloscope – something that can be stimulated and then measured, environment independent. However, emotions are not tied to the product or service themselves, but to the interaction between the item of study and the internal and external context. A lost tourist trying to catch a train has a different experience with Google Maps than a couple looking for a place to dine on a romantic evening. A college student shopping for her first laptop is different than a business executive who needs a new laptop for a key presentation tomorrow. And both of these experiences differ from a laboratory participant being paid $200 to browse for a hypothetical, unneeded laptop. Psychophysiology allows researchers to precisely measure moments of high arousal. However, that precision is wasted if the experience observed does not accurately portray the experience in question.

In one hotel experience, Design Continuum built a hotel lobby out of foam board to get the feeling of the space (Bueno & Podolsky, 2010). This prototype worked well to develop a basic understanding for the layout and note how ideas would interact. If we were to measure the emotion of a subject pretending to be a guest in the foam board hotel, their experience would be wildly different than a real guest's. A visitor arriving from a 6-hour flight likely may not pay attention to the wall decorations or furniture at all. In fact, their motivation may be focused entirely on finding a place to sit and relax or expediting check in. Collecting skin conductance data in this foam board hotel would be rather fruitless, as the individual did not need to actually check into a hotel nor could she interact with the environment in a typical way. In order to test the experience of checking into the hotel, Design Continuum could set this prototype up next door to an actual hotel, and ask real guests to check into the prototyping space. We would then be able to observe a much closer guest experience.

Early on, when I measured the skin conductance of children in occupational therapy, I was fortunately forced to observe real experiences. The therapist would not allow me to interfere with the therapy, fearing it might upset the parents or make the therapists less effective. I hid in a corner with my mini laptop and collected data and notes. By the end of some sessions, children forgot to take off their bands and forgot that I was present. The emotional experiences observed in therapy were real, which was ideal for understanding the role of therapy on children's arousal.

In my second project, I did not value the real world experience enough. I was observing the emotional challenges of LEGO Serious Play, which was being redesigned for classrooms where 5 to 6 children would brainstorm ideas through building. The LEGO group wanted to know how to make this product more engaging. I set up a fake classroom in the MIT Media Lab. The room was surrounded with glass walls with a researcher standing in the corner taking notes. I recruited five individuals who knew little about each other, and instructed them to use LEGO Serious Play, without setting any expectations. In this makeshift setting, participants were anxious around one another. They treated LEGO Serious Play questions as part of the research questionnaire, and had no reason to put effort or focus on the questions. There was no teacher telling the children what to do. I learned the consequences, when 30 minutes into one study all the boys started building spaceships and attacking each other's models, disregarding the assigned task, which correspondingly created many skin conductance responses. This mock classroom still provided insights, but I was unsure if the anxiety or boredom I measured would occur with LEGO Serious Play in an actual classroom. We were unable to measure the stress of a teacher with a tight deadline giving instructions to a classroom of 30 children. In order to observe the experience of LEGO Serious Play, I should have brought LEGO Serious Play to real classrooms.

For the next study, Hasbro wanted to understand people's emotional experience playing a board game (Hedman, 2011). I planned on providing children Clue and then watching the families try to teach the game. As I began to design this research project, I was taking multiple ethnography classes. One of my professors, Dr. Graham Jones, challenged this approach. How does the child's perspective change when told, "Here's a game, now play it"? The child is no longer "playing" a game but rather performing a task for his parents and an MIT researcher. Taking his advice, I carried a large duffle bag packed with 12 different board games, and let children choose any game they wanted. With this switch, children were able to dictate the course of play, much like a child asking to play a game after dinner. This alteration helped me discover a key insight: the most exciting part of a board game is opening the game up, like unwrapping a present during the holidays. And right after opening the box, the most boring section occurs, reading instructions, draining children's enthusiasm. The head instruction designer still uses this insight in his talks (Goodfellow, 2012).

As I began to embrace this idea of creating real-world experiences, new challenges started appearing. When working with Best Buy and Lowe's, finding people who wanted to buy the products was essential. With Best Buy, my assistant stood at the front door and asked everyone who was coming in what they were looking for. If they said they were interested in the section we were studying, we asked if they would like to participate in the study. This method enabled us to observe real emotional states – people's confusion, need for special help, and meeting or not meeting specific expectations with which they had come into the store. However, this approach biased our participant pool to those who had time. Interviews were rushed as people did not have hours to spend on an unstructured interview. I remember one man who told us his kids were in the car as we began to wire sensors on him.

For shopping, one alternative is to recruit people via company friends or craigslist. With this method, people need to have the intention of buying the item you are observing. I recommend calling interested people and asking them to describe in detail why they want the products they mentioned. If they do not have interest, observing emotional responses becomes much more difficult. What is emotional about finding the right blender if participants do not need one? With retail specifically, I would ask participants to buy an item at the beginning and then at the end inform them that buying is optional.

With the New World Symphony, we had a similar challenge observing the right type of people with the right type of experience. We wanted people who signed up for a concert, but we also most wanted people attending a concert for the first time, novice attendees. Consequently, few people attending were true novices, so we were unable to gather much data on the real emotional experience of a novice. The most helpful participant I measured was a friend of mine who had never been to a concert. A participant did not show, and I encouraged her to join the study. Her reactions were the strongest during transitions and when the emcee presented, the two spots we theorized new people would react to. However, having a mix of attendees was also valuable. We found some attendees expected and paid for a normal concert and were frustrated that the "Symphony with a Splash" presented a shorter, more interrupted program. If we had only observed participants like my friend, we would have missed this insight.

In addition to finding the right participants, researchers need to setup the right motivations. In the second LEGO study, we were hoping to measure children as they played with a difficult LEGO Technic set. The LEGO Group told every child participant that people from LEGO headquarters were going to observe how they play with LEGO sets. Many of the children were overly excited on my arrival, running to the door to greet me and look at what LEGO gifts I might be carrying.

In order to reduce the excitement of the LEGO study, I spent considerable energy downplaying the LEGO headquarter aspect. I told the kids the project was a sensor study and I was only interested in seeing sensor data of them as they sat still. Children sat at the table bored for about a minute or two. A few children asked to play with computers, which I used as an opportunity to observe how children react to games compared to LEGO sets. For the children playing computer games, a few minutes in, I would let them know the computer did not work with our sensors, and they would have to put it away. Fortunately, I had some LEGO bricks in the back of my car, if they would like something to play with. With this approach, children opted in to playing with our test sets and also could play as they like.

In this LEGO study, we were interested in how children played with a tablet version of the instructions. Part way through building, I would suggest that I had some tablet instructions that I forgot about, in case they might be interested. I needed the children to believe the tablet was optional and not necessary. With this little tweak, I could see children behave in a more realistic way. One child turned off the tablet within 10 minutes because it was too frustrating to use. Other children opted to use the paper instructions from the beginning. Only after the interview, would I let children know of our intent to build a new app. Compare these more natural interactions to standard research. In a previous test ran at a local school with many participants simultaneously building, children reported the tablet was more fun than typical instructions, and no one turned off the tablet instructions. Different settings will influence participants' affect and corresponding behavior, and researchers should consider carefully how the environment will affect participants.

3.3 PLACEMENT OF ELECTRODES

Traditionally, to measure electrodermal activity, electrodes are placed on the fingers or palms, as these locations have a high concentration of sweat glands. When possible, I recommend to continue this method and measure skin conductance on the fingers, though I use the middle phalanx to keep the electrode more firmly placed. I have tried the palm, but have difficulty keeping it taped down when participants move their hands (which they do often). There are times when the middle phalanx is not optimal. If subjects will be moving their fingers, the electrodes are much more likely to become loose or show artifacts with changes in pressure. I have had problems measuring from the fingers when participants type on keyboards, write, play with toys, and grip a steering wheel. Additionally, children have tiny fingers, and the electrodes have been hard to keep on their fingers. With children, I moved the sensor from the fingers to the arch of the foot, another location with dense sweat gland distribution. If people are walking or moving their feet substantially, I try to avoid the foot.

When measuring people with special needs like Autism Spectrum Disorder and Sensory Processing Disorder, having sensors on the hands can prove too distracting for the individual. In these cases, I measured above the ankle (Hedman et al., 2012). While not ideal, the children were able to wear the sensor over multiple days. Overall, I recommend researchers try measuring from multiple sites in the beginning and chose the location that generates the least amount of noise while still having a strong enough signal.

3.4 OBSERVATIONS

3.4.1 How to Observe

When I first started with the skin conductance sensors, I had little appreciation for observations. Initially, the occupational therapy setting was only necessary as a place to collect sensor data. As the data was being recorded, I would start noticing events that were happening. “Oh, wow, there is a jump when that child is climbing a wall, that’s interesting.” Not thinking much of it, I excitedly presented the occupational therapy staff a graph of multiple sessions so far, with responses changing across time. “Look, the child’s response is changing!”

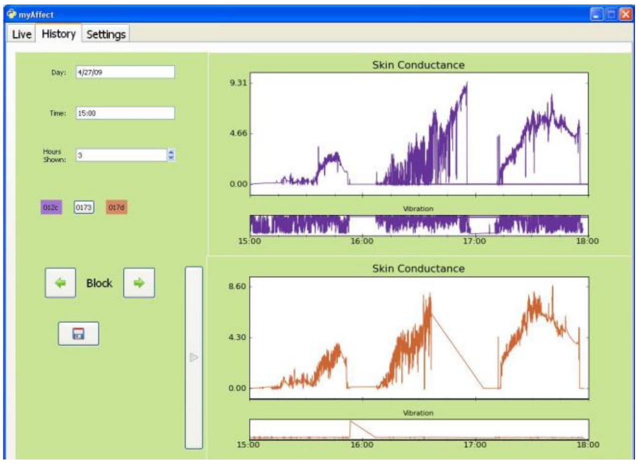


Figure 3.1 Initial Interface for Measuring Children during Occupational Therapy. Orange and Purple are skin conductance responses from two different legs.

I was more interested in the sensors being able to measure than what they measured. That perspective changed when the therapists asked why the responses occurred. When I brought the results back to my advisor, she asked the same question, “What was occurring during these moments?”

In my observational methods class at Harvard, I learned that sitting and watching people, even without sensors, is a strong way to learn about experiences. I started closely watching people while I collected their skin conductance. I observed children rock back and forth as their mother explained board game instructions. I observed people miss entire selections of items in retail spaces as too much information was presented, and I listened to what people said to one another. Often times what is going on is so complex that I re-watch the scenario multiple times.

When conducting thick psychophysiology, videotaping is a necessity. I can spend two to three hours replaying one child's 20-minute interaction with a LEGO set. Take for example Ryan's skin conductance as his father helps him build a LEGO Technic set. I see his skin conductance decrease as he interacts with his father, but then responds suddenly.

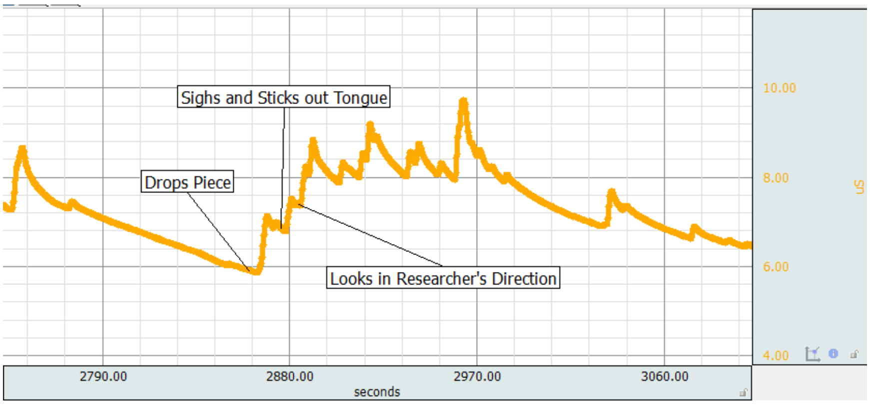


Figure 3.2 Reviewing video provides context for why a skin conductance response may have occurred.

Only by replaying the video along with the skin conductance can I dissect why Ryan's response may have changed. Initially, Ryan drops a piece and is likely reacting to the dropping. The father picks up the piece and takes the model from Ryan and starts building himself. Ryan sighs and sticks out his tongue. The third peak occurs when Ryan looks over at the researchers, perhaps in embarrassment. From close inspection of the video, I do not simply label the response as Ryan dropping a piece but a possibility of multiple factors. The video allows researchers to identify the complex interactions that occur around a skin conductance response.

While I have not been able to do this yet, I am excited to add eye tracking to future observations as well. When children were reading LEGO instructions, I would see a reaction while they are looking at the instructions and wonder whether they were looking at the page number, a complex piece, or a future step. Similarly, in shopping, I would like to know what part of signs create responses, as I cannot be sure what exactly people are reading. In the New World Symphony study, I would wonder what people were focusing on visually. One participant reacted in an unusual spot. When I replayed that section, he told me he was watching the triangle player the entire time, and that player had finally started playing.

3.4.2 Observations Affecting Participants

I am often asked if putting a sensor on people makes them self-conscious, which in turn alters behavior and responses. To prevent self-conscious behavior, I present the sensors in a non-emotional context. “The sensors will help us note times when you are interested, but I still will not be sure what you are interested in until you tell me in the interview.” With that description, most people do not report feeling self-conscious or embarrassed.

However, a much stronger effect seems to come from video cameras and human observation. Recording a participant can make them anxious. During the board game study, one mother from Jamaica was playing Connect Four with her son who spends a significant time at the Boys and Girls club. The parent and child started arguing more and more about the rules, likely because the mother did not read the directions. She accused her son of cheating, and he started laughing. Eventually her son goes to the bathroom and she looks up at the camera. The moment she looks up at the camera is the largest skin conductance response during the entire session. I interpreted this as a moment when the mother felt vulnerable, perhaps judged, as she was not performing in the way society expects a mother to perform. I saw similar responses from children as they were asking their parents for help in front of the camera or looking at the camera lens.

Cameras can significantly alter an emotional experience. There is a sense of judgment and performance if someone is being recorded. I know some researchers have participants carry the cameras themselves, but this likely will alter the experience as well. Alternatively, people could wear cameras so they are not being observed and still allow others to see what they observe. An ideal situation would be to tell the participant they may be videotaped, and then to hide the cameras from easy view. In public places, I then place a sign letting the public know videotaping may be in progress.

3.5 INTERVIEWS

Observations provide a limited perspective of a person’s experience. To complement observations, I conduct long, unstructured interviews. In the occupational therapy study, after 70 hours of video, I had a consistent question: what were the therapists intending as they conducted therapy? I wish I spent even just two minutes after each session asking the therapist how the session went. In the Hasbro study, I observed one mother who had a clear skin conductance level decrease as she read the board game directions.

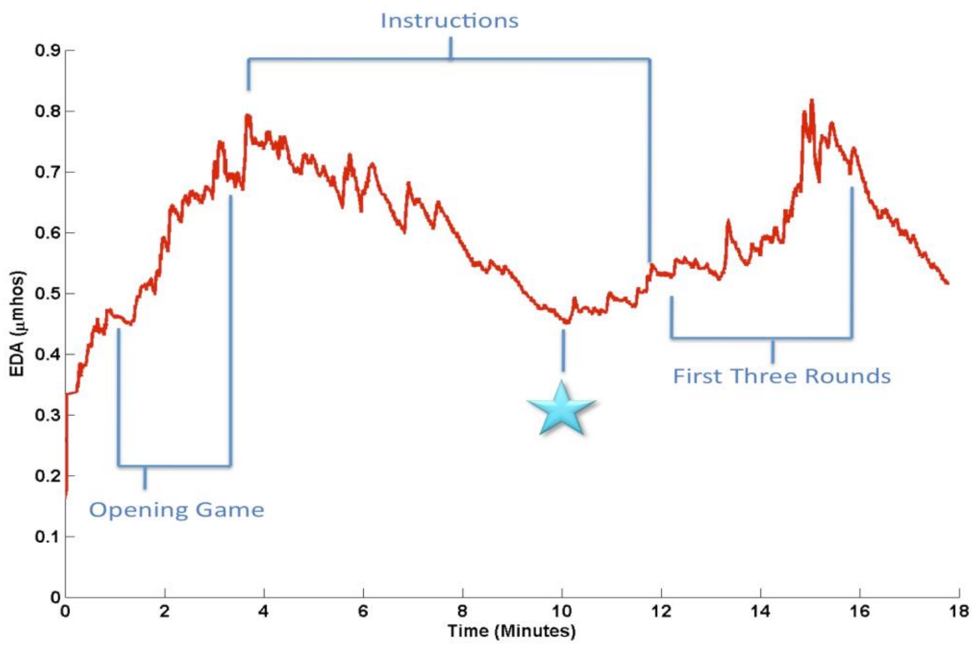


Figure 3.3 A mother’s skin conductance decreases as she reads the instructions.

Her children were starting to rock and ask if they could play, so I assumed that she was bored too. Much to my surprise, in the interview without the psychophysiology showing, the mother tells me she felt calm reading, like reading a newspaper. Interviews are essential to help us understand participant’s perspectives and personal interpretation of their experience.

When I videotaped groups as they played with LEGO Serious Play, one woman reliably produced large skin conductance responses before she started to speak to the group. This woman’s skin conductance decreased after she began talking.

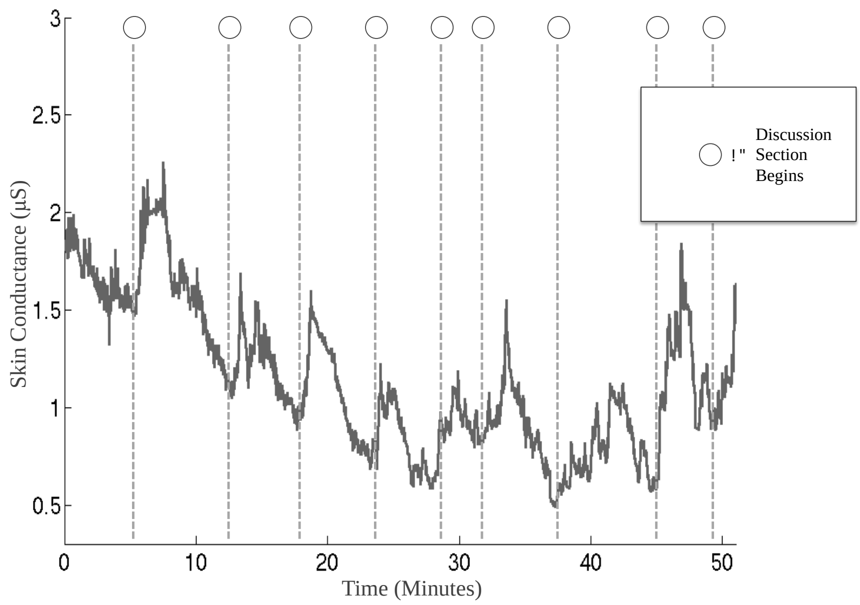


Figure 3.4. Stephanie's skin conductance increased before she presented her LEGO model.

Compare this response to another woman who reacted in the exact opposite way, her skin conductance decreased, and then increased when she talked. I needed interviews to help me understand what the two participants were thinking and why these two responses differed. The first young woman, whose arousal increased until she started talking, reported she was a bit anxious about speaking. Specifically, she did not know the people in the group and her model was not complete, so she was worried how she would talk about an unfinished model. The second woman, with the inverse response told me how emotionally important her family was to her, and speaking of a beach trip brought back an endearing memory. Interviews helped me understand a person's perspective and why an emotional response may have occurred.

3.5.1 Conduct Interviews as Soon As Possible

In terms of identifying emotions, interviews need to be done immediately after a person had an experience. I interviewed some therapists a year after they had conducted a therapy session, once the data was analyzed. This method does not work. Our memory of emotional events are subject to forgetting and become biased over time (Levine & Pizarro, 2004; Levine et al., 2009), and therapists were unsure what their intentions were or whether the therapy was effective, even with video to review. In a shopping experiment, I had participants tell me about their purchase after we drove back to their home. This gap in time was enough to prevent people from remembering the particular details, "Well, it was a pretty normal experience” or “I don’t recall much.” After a few of these interviews, I now conducted retail interviews in the store space. This added the benefit of allowing people to reflect on the display, layout, and messaging. However, when talking about employee interactions, I try to find an empty room, as you want participants to have the freedom to express problems without being judged.

3.5.2 Use Open-Ended Questions

The worst question one can ask is, “Were you stressed when you _______”. “Yes.” or “No” tells little about a person’s perspective. My general first question is, “So, how did it go?” Perhaps when watching the concert, someone really liked the visuals. In general the research should use open-ended questions. People tend to be much more descriptive when they are telling me stories of what happened. I have participants walk me through what happened in detail. As they describe events, I ask them to describe even more what they were thinking or feeling. After a detailed interview, I then look at the skin conductance, and ask them about any events they may have missed in the recollection.

3.5.3 Ask People about Exact Moments

When possible, I also play back the video at moments when responses occurred, telling participants that some highlighted events might be emotional while others not. Re-watching video can remind people of all the events that were occurring at a moment. People’s recollection of specific events also provides another measure of how salient an event was likely to be. If a participant does not remember a strong emotional experience at a moment, that emotional reaction may not be strong. When interviewing people about specific sections, I still have participants describe the event in a story fashion. I remember asking engineers how they felt during moments of LEGO Serious Play, and their answers were often, “I played with the bricks”. “But how did you feel?” “I played with the bricks.” I tend to not ask people how they feel anymore. Rather, I ask them to describe what happened. Describing events tends to bring out emotional descriptions without prompting.

3.5.4 Do Not Show Skin Conductance

Do not show people their skin conductance during interviews. Skin conductance is an easy way for people to suggest they were supposed to be feeling something, likely stress, and emotions are easy to implant into participant’s memories. Showing someone his or her skin conductance is like asking the question, “From my expert view, I believe you were stressed at this point. Why do you think that was?” You want to know if people do or do not recall anything, as this helps show that the moment likely was not as arousing as the skin conductance might initially suggest (LaBar & Phelps, 1998). There are some special circumstances where showing skin conductance is appropriate, like teaching children about emotions or demonstrating the sensor technology, but for design research, the sensor data typically should not be shown to participants until after the interview.

3.5.5 When Interviews Do Not Tell the Whole Story

While one person drove in a difficult setting, I measured his skin conductance. His knuckles were clenched on the car's wheel. His wife said, "You sure look stressed, are you sure you are all right?" I asked the driver what happened at that moment, and he reported feeling no stress. This example brings up epistemological questions about what is an emotion. I interpret this data in two ways. First, from a memory perspective, this event was not stressful. He will not be telling his friends how stressed he was. However, consequences of a stress reaction were also probably present. He was probably more easily agitated, focused, and tense. I would expect that after the drive he reported feeling surprisingly aroused but attributed the feeling to another issue. When reported emotions and physiology differ, I tend to describe both the report and other observations that tell a different story.

Occasionally, a specific moment will seldom be remembered as arousing. When asking a symphony attendee what the most arousing part of a concert was, few people will mention the transition between songs. Transitions are not something we think about when reflecting on emotions in music. However, the collected data showed consistently large skin conductance responses during transitions, even when participants do not clap. Perhaps, in a concert setting we expect the medium to be the cause of emotional stimuli as it is our intended focus. From a grounded theory perspective, I know that transitions, loud noises, and talking all create skin conductance responses, and so I was not surprised to find responses during transitions. Despite no verbal confirmation, I concluded that transitions help arouse audience members.

Interviews with children also tend to produce unreliable reports. I remember Luke in the board game study explained how he read the rules to his mother, despite him not knowing how to read more than a few small words. Mason told us Technic was his favorite set and he would recommend it to his friends. When I watched Mason on the video, I saw him stop building and ask his mother if he can play something else. Mason's mother told me that he almost never builds by himself, which probably created anxiety – something Mason did not describe directly. In all of the above cases, I had to rely on observations and other sources of information to help me characterize the skin conductance responses.

Other populations like the elderly or people with alexithymia may prove difficult to uncover emotional experiences from interviews alone. These extreme cases accentuate the larger question about why we privilege self-reported emotions as a gold standard when self-reports themselves can often be unreliable.

I was hesitant to put these interview challenges in my dissertation, as self-report being inaccurate is usually assumed to be the exception to the rule rather than the norm. Most often, if people do not remember an emotional event occurring, I tend to not put heavy emphasis on that skin conductance data, while respecting that the response may be emotional and the participant is unable to recall or talk about it. Random or artifact skin conductance responses are so frequent, that I believe people's reports are more accurate than the skin conductance alone. From a design perspective, I bookmark these non-confirmed responses as an interesting event and interview other people about that moment to see if other people do recall these events as emotional. This is not to say that skin conductance is dependent on self-report, but more to say skin conductance combined with self-report is more accurate.

3.6 ANALYSIS

3.6.1 The Generalizability and Unique Emotion Problems

With my occupational therapy work, I took 70 hours of skin conductance with video and tried to correlate responses to activities: do children in ball pits calm down easier than children on swings (Hedman et al., 2012)? One statistics professor on my master's thesis committee encouraged me to conduct a strong, statistical analysis. I recruited undergraduates to painstakingly go through all 70 hours of video and label the moments when children changed activities. I labeled all of the videos for body position, activity, and conducted more data processing than I ever had hoped to do. All of that work produced one small result: children are more likely to increase skin conductance when they push themselves on a scooter board, $p<0.064$ (Hedman et al., 2012), and this result was probably not significant with more than 8 tests. Furthermore, scooter boards involved significant physical activity, so we could not assume the difference was emotional. So, did children's skin conductance not change with therapy, despite therapist's insistence that therapy was about helping children regulate their arousal? Actually, the skin conductance responses were notably dynamic. I saw one child become stressed seeing a robotic toy, a child begin to calm down while she painted, and multiple children become anxious at having to talk into a microphone. But traditional statistical analysis skipped right over these findings in an attempt to find a generalized trend.

Real-world emotions are dependent on an individual's personal experiences. One child we worked with in occupational therapy was overly excited as he entered the ball pit. He would start announcing, "Ball pit, ball pit," at the beginning of each session. Not surprisingly, his skin conductance increased considerably while in the ball pit. Compare this to another child's response to the ball pit. Her skin conductance dipped down as soon as she entered; the ball pit was relaxing for her. If the two children's responses are combined, a rather flat, highly variable response is produced, which is far from the truth of either child's experience. From a design perspective, we want to understand individual difference, and this flat response is detrimental. If 70% of customers love your service and 30% hate it, you have a problem with that 30% and that should not wash out with the average.

A secondary problem is the complexity of emotional experiences. A lab has highly controlled stimuli, but the real world has constantly changing and multi-layered stimuli. When Anna lies prone in the ball pit, the ball pit is not the only factor influencing Anna. The therapist tells Anna it is time to be like Eeyore, time for calming down. No children are in the room. Anna is interacting with small stuffed animals found in the ball pit, on which she focuses intently. Anna is partially covered in balls, lying on her back. The lights are low. This list goes on. With so many factors, Anna will likely never be in the ball pit with the exact same internal and external context as before. This unique experience problem persists with typical adults. When shopping for a blender is the shopper in a hurry? Is her child in the car? Is her blender broken and she needs one that afternoon or is she browsing? Does she learn visually or through interacting? Did she look up blenders on the Internet beforehand? This list also goes on. Statistical analysis wants to take this long list of factors and conclude buying a blender is either stressful or not, but in truth many different aspects (more than there are subjects) influence an emotional experience. Because emotions are complex and individuals experience the world differently, I focus on individual responses and describe specific reactions to products and services.

3.6.2 Analyze Right after Data Collection

After two months of data collection at Google, I analyzed people's skin conductance responses when all data was collected. Unfortunately, I found one hundred new questions that I was not thinking about that I wished I had started noting from the beginning. I also learned too late that some of the sensors were not working well while driving: the sensors needed more gel for proper recording.

Compare my Google analysis to my more recent work with LEGO instructions on the tablet. I went over the data immediately after collection. After analyzing the first child's data, I learned my presence had a large effect on the child's experience and skin conductance (he wanted to impress me). In the subsequent studies, I removed other researchers and myself from the room. We also saw a large number of sensor artifacts in the sensor on the hand, so I moved both sensors to the arches of the feet. We initially asked the mothers to not participate so we could observe how children cope with stress. During a fluke interaction, the first mother started helping her child, and the skin conductance changed considerably. By analyzing data that day, we learned that mothers could play an important role in helping children play with challenging LEGO sets. Consequently, we allowed participants to choose how they interact with parents in all future sessions. Researchers should put time into their daily routine to analyze the data as it is being collected.

3.6.3 Combining Video and Software for Analysis

I remember in my early days, the skin conductance response and video were separated, and I'd be curious what happened at a moment and it would take 5 to 10 minutes to pinpoint that moment and watch that clip. If there are 100 skin conductance responses during a session, video analysis becomes daunting and unattractive. Software-wise, the skin conductance needs to be tied to the video directly. If I click at a base of a peak, I need to be able to see exactly what occurs at that moment, instantly.

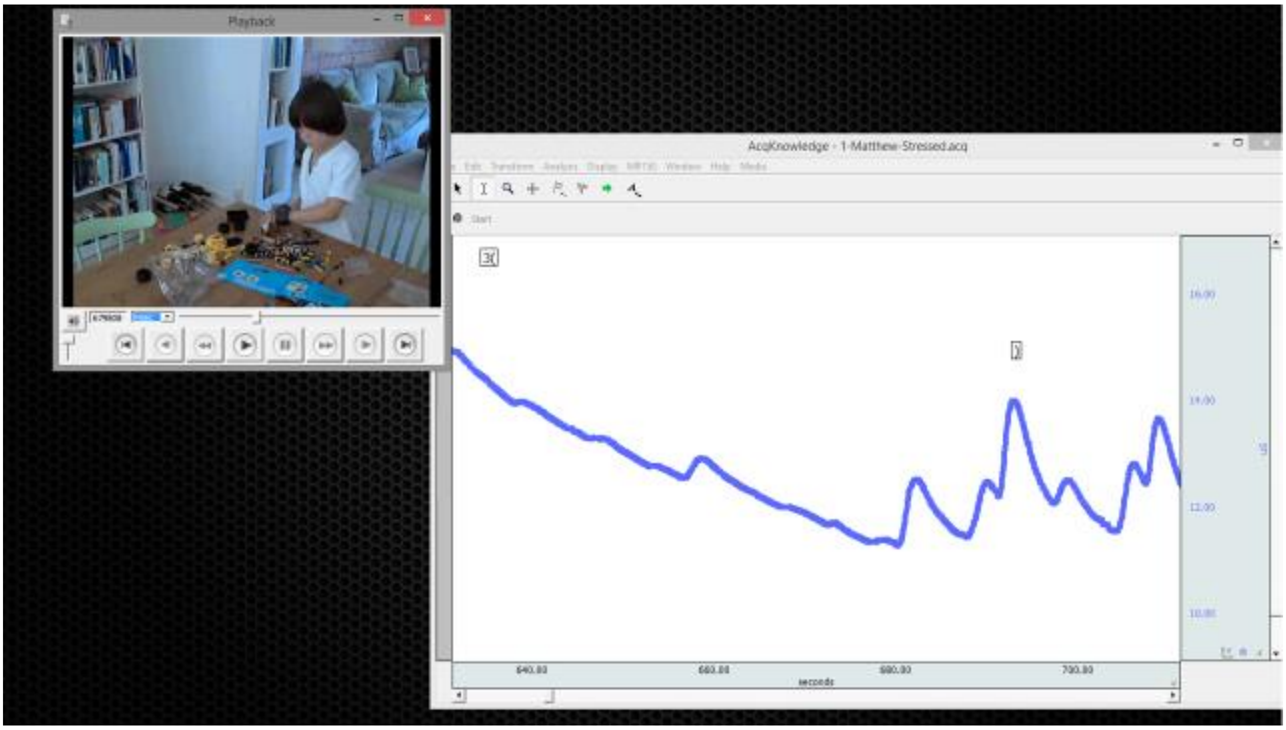


Figure 3.5 Interface used for combining video and psychophysiology.

Videos need to be immediately available after an experiment. Having to render videos or graphs can take hours and researchers will naturally fall behind. As the number of videos that need to be reviewed increases, the chances of the researcher never viewing the data also increases. Sometimes a client will have a question, and it is great if the researcher can pull up the file in a few seconds to tell that story. Consequently, researchers should not use the highest definition video as the number of files that can be stored will be limited, and access will take too long (I imagine in 10 years, a graduate student will laugh at that last statement).

3.6.4 Selecting Skin Conductance Responses for Analysis

Once the video and skin conductance are properly aligned, I zoom out and see the entire Skin conductance across the session as a whole. Are there clear places where responses appear to change? Are there trends up or down? Artifacts can make skin conductance appear flat, so researchers may need to zoom around these spots or filter out those noisy jumps, before viewing. With the Biopac software, I set all values below a threshold or above another threshold to the mean skin conductance response. In Matlab, I set the values to NaN so they do not appear on the graph.

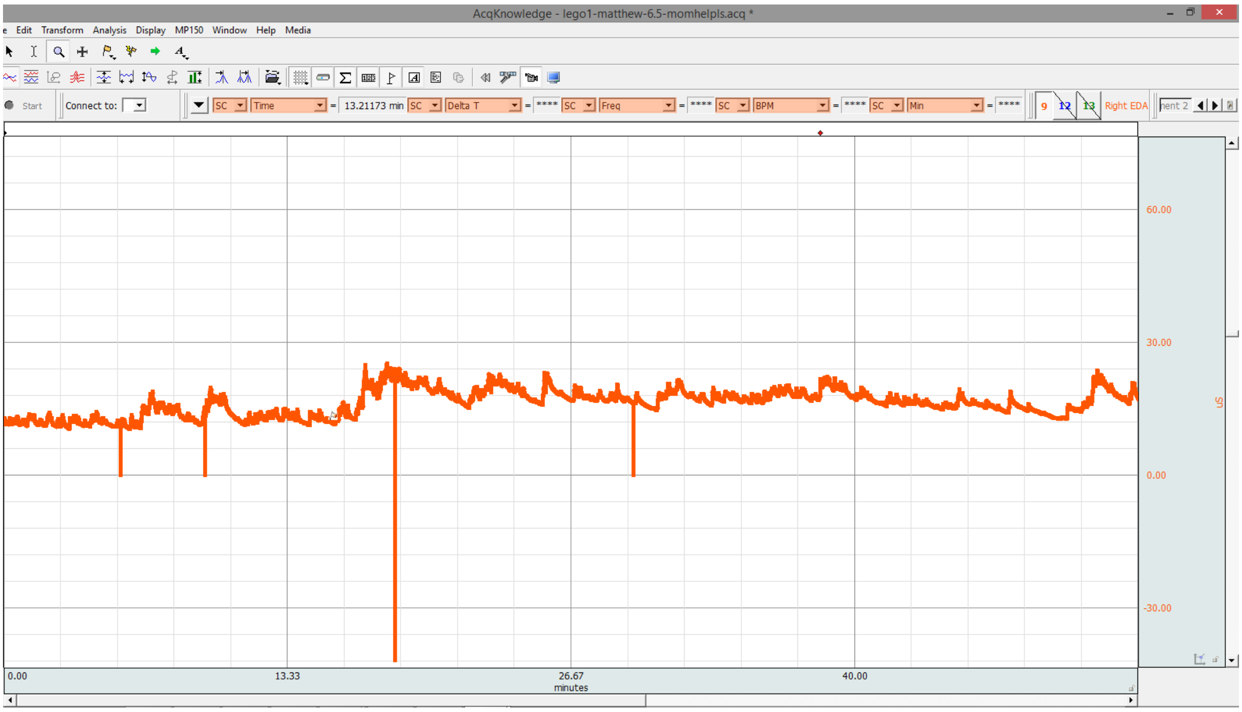


Figure 3.6 Signal dropouts make the skin conductance responses difficult to observe by affecting the auto-scale and should be removed before analysis.

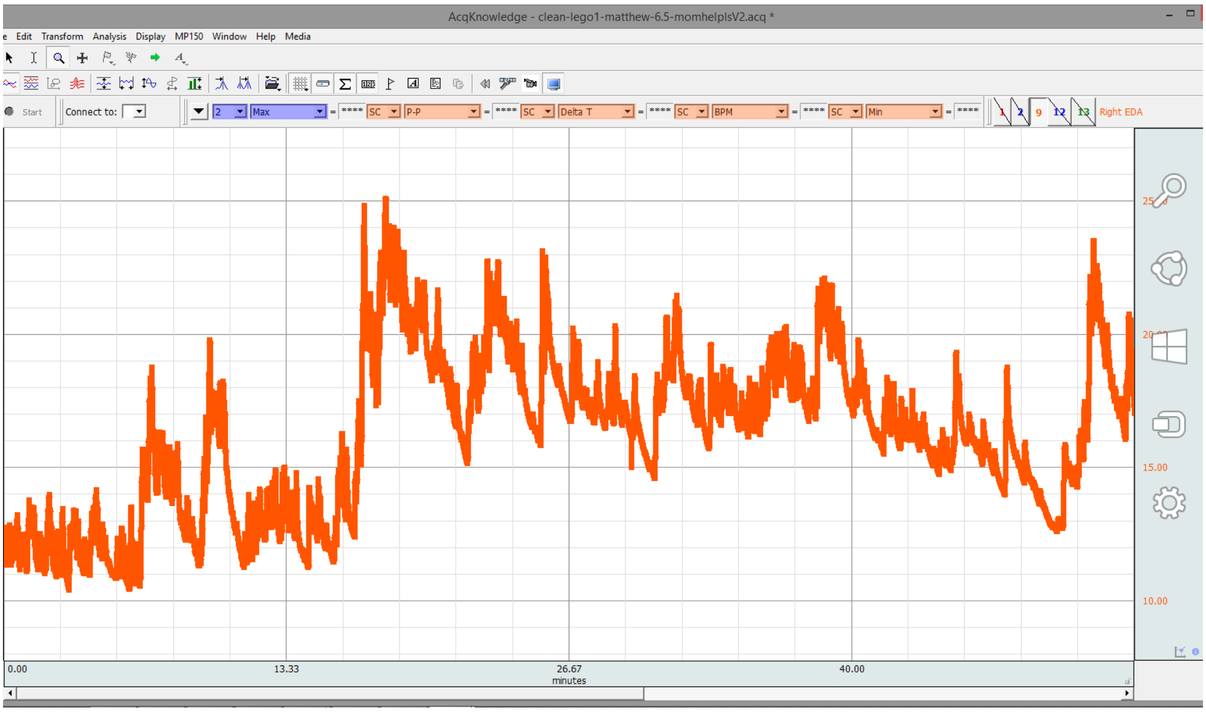


Figure 3.7 Same response as above, with dropped signals removed.

The researcher will need to determine what level of granularity is appropriate for analysis. One could re-watch the video and attempt to label every response, but during long sessions there will likely be over one hundred responses, which will take many hours. Additionally, many of these responses are non-specific; digging into the trenches of every skin conductance response can be dangerous, as the small responses likely are physiological noise, and could be incorrectly attributed to unrelated events. If the skin conductance response is less than 0.05 uS, I do not include it in my interpretation at all.

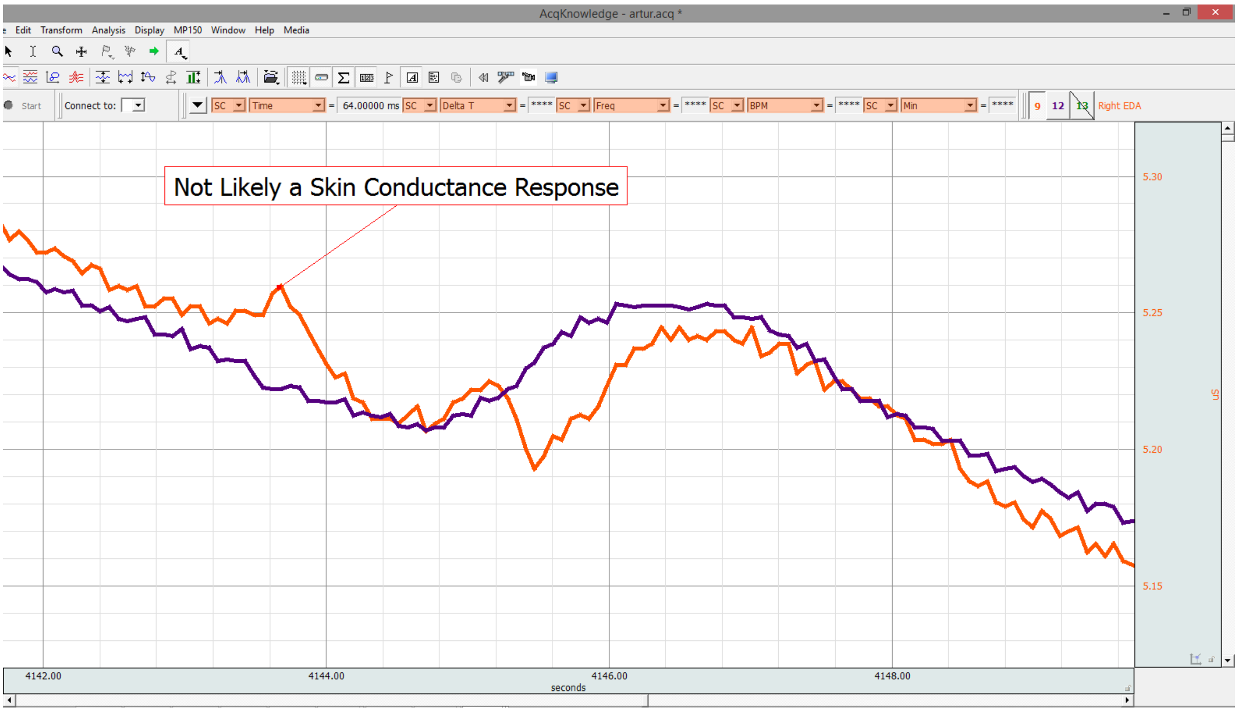


Figure 3.8 Too small of responses may be artifacts.
Here skin conductance is measured from the arch of the left and right foot. The orange peak noted shows changes that likely came from movement as identified by the video.

I tend to rank the responses by amplitude. I then pick a threshold and analyze skin conductance responses above that threshold. One way of determining a threshold is to only look at the top 20 skin conductance responses, as time for analysis tends to be the most limiting factor. Keep in mind that some of the top responses can come from large movement artifacts, so researchers will need to extend their list of analysis beyond the initial responses.

Here is one woman's skin conductance during a classical concert. As can be seen, there are moments with much larger responses.

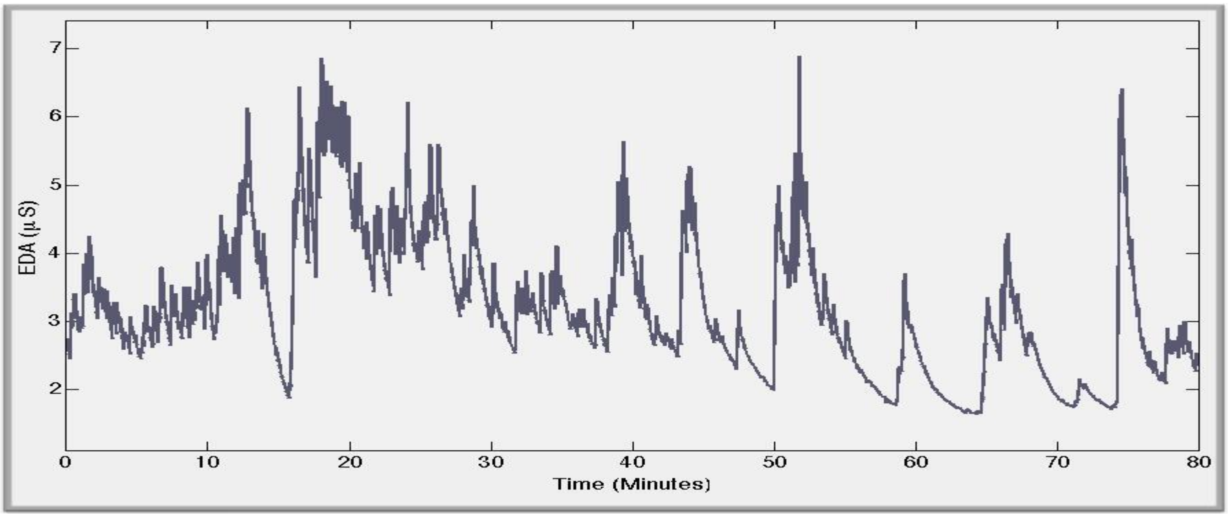


Figure 3.9 Anita’s skin conductance from watching the New World Symphony.
Some skin conductance responses are larger than the average.

In the next plot Anita’s skin conductance response amplitude is plotted over time. By only analyzing the highest responses, we reduce the amount of time needed to analyze, and reduce the probability of measuring a non-specific response, and increase the likelihood of observing a highly salient event. Many of these high responses can be due to large movements, so a larger sample size of responses may be needed.

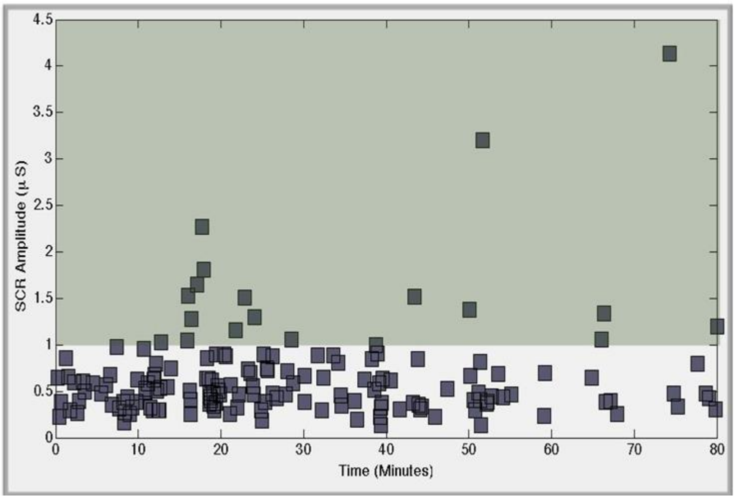


Figure 3.10 Plot of the amplitude of skin conductance responses.
The box shows a theoretical filter that only analyzes the 20 largest responses.

While pre-selection logically makes sense, as a designer, I often find myself homing in to important points more closely as well. For example, the moment Mason quits building Technic and asks his mother if he can do something else is a critical moment from an experience standpoint. I scrutinize all of the responses, large and small, around the time

Mason quits, in order to determine what might have been emotionally provocative at that moment. Interestingly, Mason's last skin conductance response before he quit was when he was looking for a LEGO piece.

Figure 3.11 The small skin conductance response as Mason looks for a piece would not catch researcher's attention if they only analyze large responses.

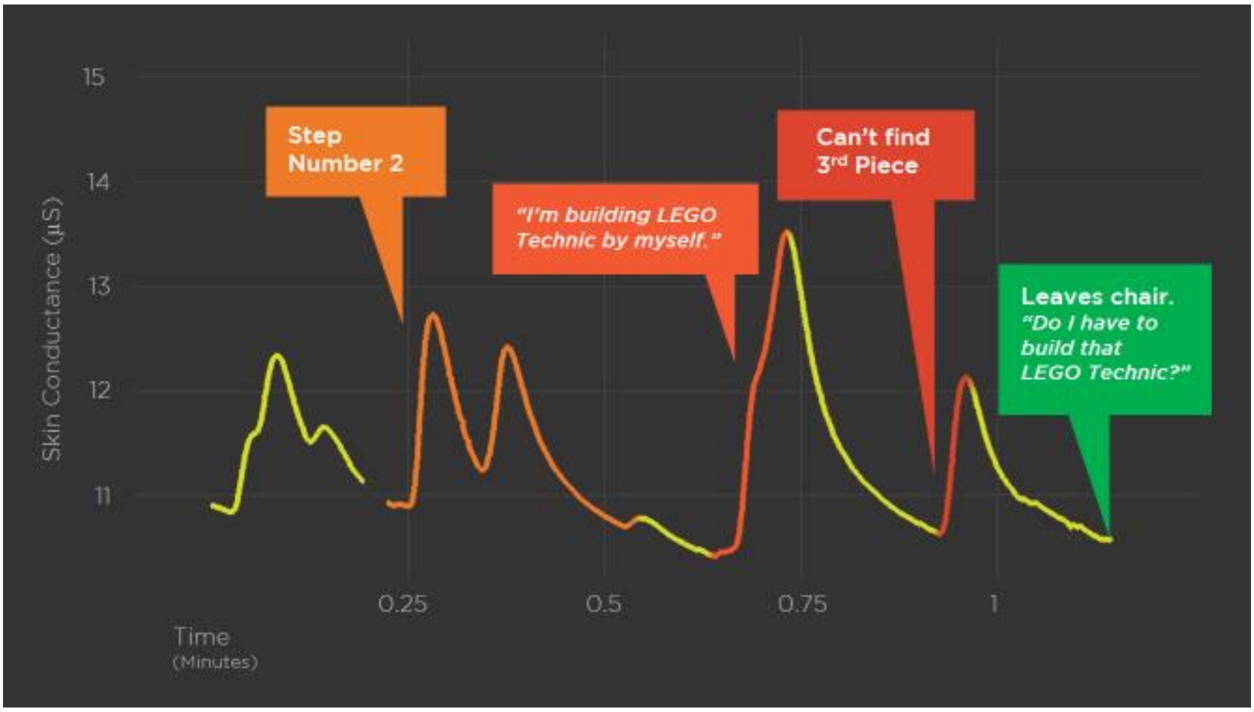


Figure 3.12 Only looking at the responses that occurred right before Mason quit helps emphasize the searching response.

3.6.5 Identifying Possible Causal Elements of a Skin Conductance Response

In order to analyze a skin conductance response, I play the video 7 to 10 seconds before the response. Understanding the context is important, so if researchers do not remember this section of the video, they may want to go back a few minutes and watch events unfold, to better orient themselves to the moment. I note all the possible changes that occur 1-5 seconds before the response (we start at 10 to be able to see changes occur). Actions that started outside of those 5 seconds or after the response occurred were not likely the cause of the response. This method reduces the number of possible causes of a response dramatically. Typically I only see one or two things occur before a skin conductance response.

The first step is to rule out movement artifacts, as these responses can be otherwise misleading. Many of the responses researchers encounter will not be due to salient events and it is imperative to account for these alternative responses. When watching, determine if any physical movement occurred. Do we see the person's back move or their fingers pressing hard, breathing hard, etc.? Below is a list of activities I have seen that can create large skin conductance responses:

Walking

Leaning In

Shifting body posture

Lying or Sitting Down

Standing Up

Moving arms or hands

Tensing muscles

Increasing or decreasing pressure on electrodes

Lifting an item

Breathing hard

Talking

These activities occur regularly in day-to-day interactions. If a response occurs while an individual is doing any of these things, researchers cannot assume the response is due to a salient event alone. This is especially true if the physical movement occurred right before the skin conductance response. However, researchers cannot conclude that psychological arousal had no effect either. When I saw children open a board game box, they produced multiple skin conductance responses. These responses were likely due to the moving of hands, but this increase could have also been due to the children being excited to open the box or due to both factors interacting. When I see a slight movement before a response, I tend to go back a bit and see if the movement was occurring beforehand. For example, if an individual is walking along a sidewalk for 5 minutes, and the first skin conductance response occurs while she is still walking, the skin conductance response is likely not due to walking. However, she could also have changed how she walked; perhaps she tensed her leg or took a heavy breath.

One way of accounting for some movement artifacts is by measuring skin conductance on both sides of the body.

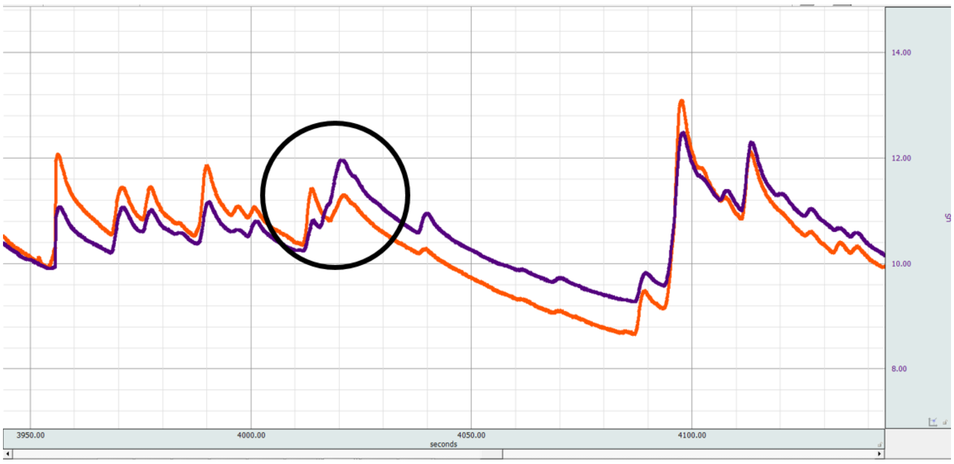


Figure 3.13 A child's left and right responses differ with movement.

In this example a child wore two sensors on the arches of his feet while playing with a LEGO set. A little after 4000 seconds (circled in Figure 3.13) we see him shift his legs, and correspondingly the skin conductance responses differ between the two signals. Nothing else different appears to be happening at this moment. This extra-large increase with the purple signal was likely due to the leg movement, so we conclude that skin conductance response was not likely due to a salient event.

Often times, multiple events can co-occur before a skin conductance response. The safest assumption is that any of the events may have created that response. If participants report a strong emotional experience with a specific interpretation, I emphasize their view more. Also, some events tend to have a higher probability of creating a skin conductance response than others. For example, a cell phone ringing will almost always create a response and so does starting to talk to another person. If I see an event occur that tends to create a skin conductance response, I will attribute the response most likely to that event. Below is an informal list of activities that tend to create a response ranked on how likely I think they are to produce a response. I rank these on both their probability to occur and the amplitude of the response. Not surprisingly, events that create large responses seem to create reliable responses as well. Please note that this is only from informal observations. This list lends suggestion for a future study that could better understand what factors influence skin conductance the most. Of course, individual differences may make some responses increase or decrease in rank as well.

High Probability of Creating a Skin Conductance Response

Physical Movement

Cellphone ringing

Preparing to talk or giving a speech

Being talked directly to

Pain

Changing velocity suddenly

Medium-High Probability of Creating a Skin Conductance Response

Talking about something important

Searching for an item

Being exposed to new information – a page turn, a start of a movie, a transition

Having a conversation

Having multiple stimuli compete for attention

Having a new sensory input – taste, smell, touch

Quickly transitioning between tasks

Medium-Small Probability of Creating a Skin Conductance Response

Having to choose between items

Being exposed to something familiar

Something abruptly ending

Talking about something important

Being confused and having to deduce meaning

3.6.6 Identifying Emotional Characteristics of a Skin Conductance Response

When Mason tells his mother, "I'm building LEGO Technic on my own," a large skin conductance response occurs. This response was likely due to Mason announcing he is building alone. However, what can we infer from this skin conductance response? Was he excited, frustrated, a combination, or not emotional? Clients consistently ask me how I determine valence, whether a response is positive or negative. From Mason's skin conductance, I was not sure about the emotional characteristics; all I could conclude from the response was that a salient event occurred.

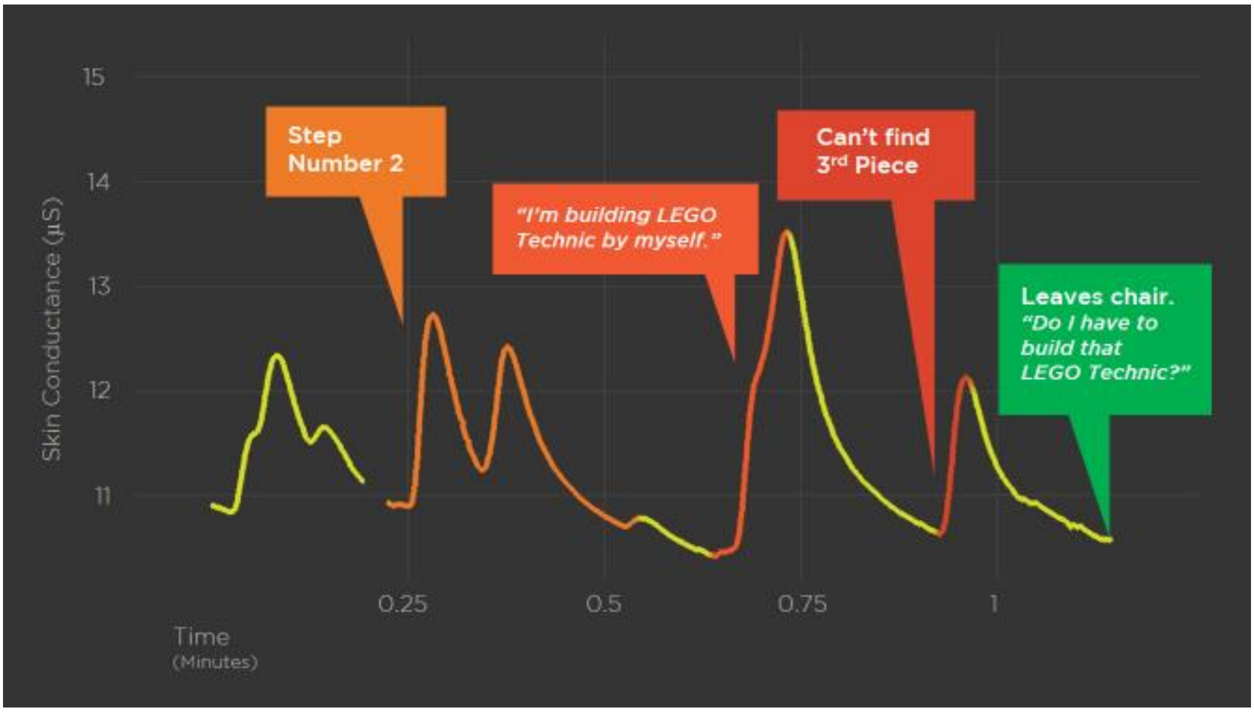


Figure 3.14 Is Mason telling his mother that he is "building LEGO Technic by myself" a positive or negative response?

I use context and interviews to help me identify the possible consequences of this increased arousal. If a shopper describes their frustration when a sales clerk offers them insurance options with the blender, I interpret these responses as being related to frustration. One participant's skin conductance increased when an audience member acted on the front stage of the Blue Man Group. The participant told me afterwards that she felt embarrassed for the audience member, but also enjoyed that section. Her report helped me characterize why this skin conductance response likely occurred. When Ryan was told he had to go back in the instructions and a large response consequently occurred, I concluded his response was negative, as he told his father he did not want to go back and asked if he could quit. When a person is about to inject themselves with a shot, I assume the response is due to anxiety, unless, of course he or she self-reports as being masochistic. The skin conductance response helps us identify moments as salient, but we must look into the context and interviews to understand why. Typically, between observations and interviews, the likely reason a salient event occurs becomes clear.

3.6.7 Absent of Responses

I'm reminded of Claude Debussy's quote, "Music is the space between the notes." Sometimes people not reacting provides insights on its own. One benefit of no reactions is there are no challenges to specificity (Cacioppo & Tassinary, 1990), if an individual did not react, they likely did not react for any of the possible reasons that responses tend to occur. I look at long periods of no reactions and label those occurrences as well. For example during the New World Symphony, I was surprised to find novice listeners not only failed to react to the music, but they also did not react to the emcee speaking. Perhaps the emcee was not the best way to reengage listeners. In the LEGO instructions study, children produced responses looking for pieces and turning pages, but when their parent built with them, the responses occurred less frequently, suggesting that the parents helped their children to stay calm, perhaps to the point of boredom.

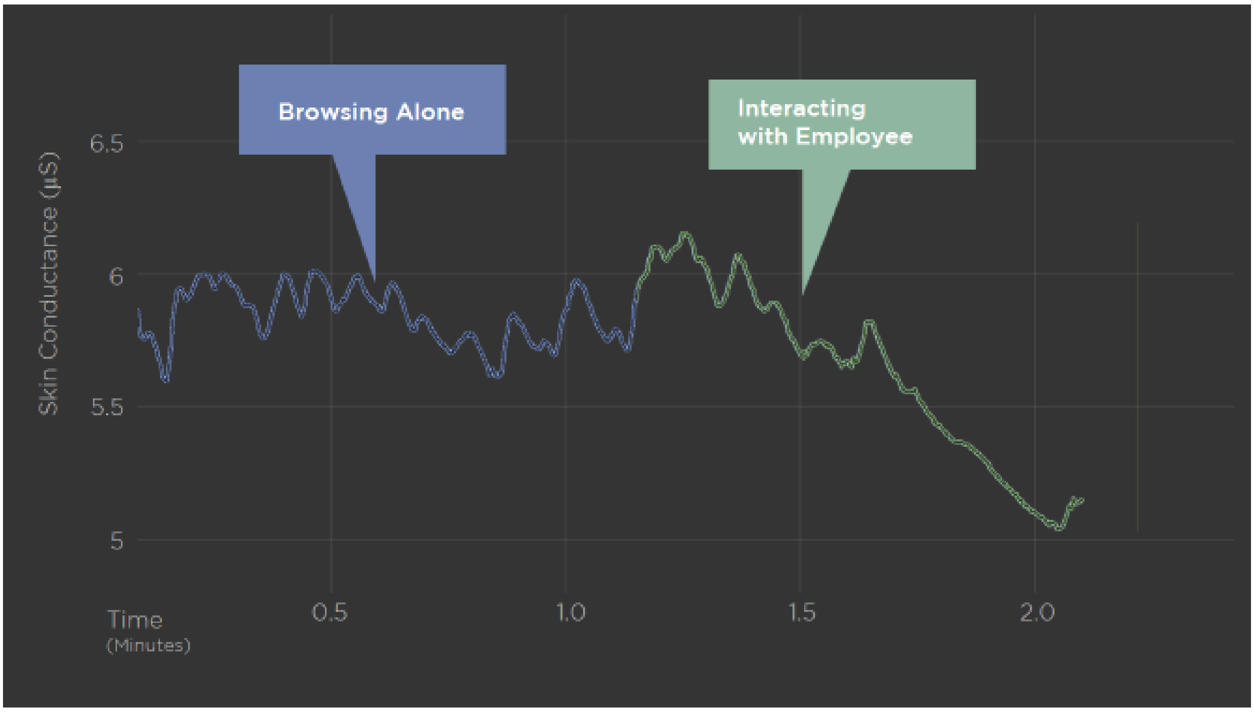


Figure 3.15 When this individual stops shopping by himself and talks to an employee we can conclude that he is probably not becoming more stressed, frustrated, or anxious.

3.6.8 The Double Response

A consistent feature during times of highly arousing events is a participant producing two or more responses so close to each other that it almost appears as one response.

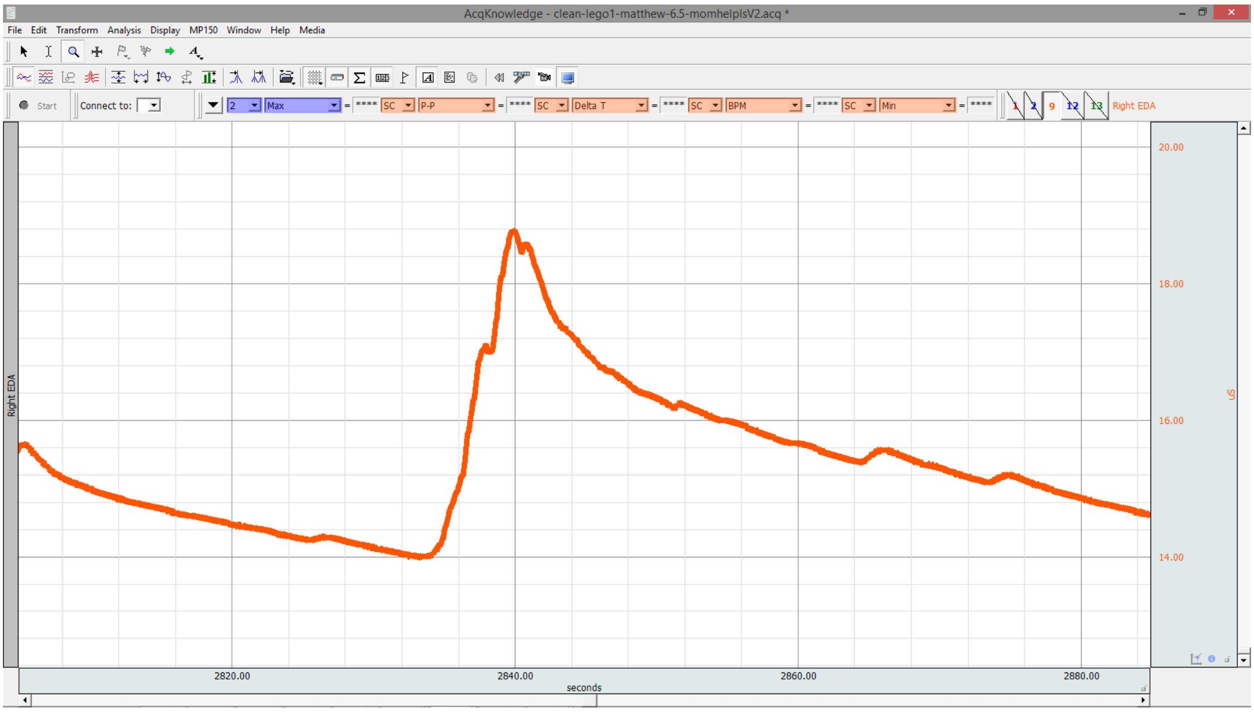


Figure 3.16 A Double Response (notice the small dip before the signal peaks).

Statistically, these double responses are difficult to interpret – do we count this as one or two responses and what amplitude makes sense? From an observational point of view, this is much easier – quick double responses are important moments. The sympathetic nervous system is so fast that the physiology does not have time to return to baseline. I typically look at double responses first, as I have found highly salient events are more likely to occur at these points.

3.6.9 Multiple Participants

When skin conductance response is measured from multiple people simultaneously, additional analysis can be made. When people watched things together, like a concert or movie, I have collected responses from the group. I have also measured multiple responses with people interacting together, like a therapist and child. For analysis, do not average participant's skin conductance levels. One person responding heavily to an event and three people not responding does not equate to a medium response. Instead, count who responded and who did not at certain events. When multiple people respond simultaneously, the stimulus was likely externally produced, making the likelihood of movement artifacts or drifting thoughts less likely. I can also start to conclude whether effects may be more generalizable with multiple, simultaneous responses.

Sometimes, the people being observed will have different reactions, in part because experiences will differ between each person. Differences in responses provide new information as well. For example, a therapist and child may react to one another's changes in arousal. In the New World Symphony, experts responded to music that novices did not. When measuring driving, measure the driver and the passengers. Being able to view multiple participants' responses simultaneously helps researchers better understand how responses differ or are similar across people within the same environment.

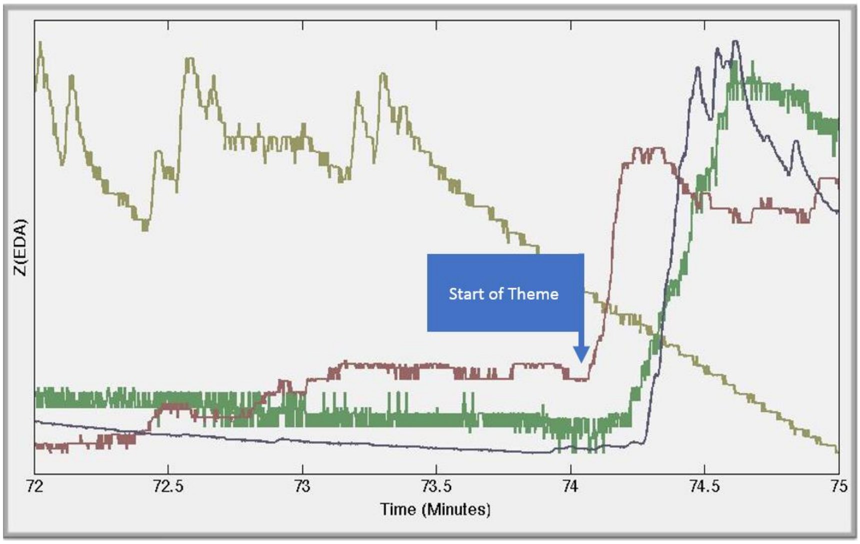


Figure 3.17 Skin conductance of four participants is measured simultaneously.

During the introduction to the love theme of *Romeo and Juliette* (around minute 74), three individuals produced a notable skin conductance response. The fourth individual already had a higher level, perhaps in anticipation.

3.7 EXAMPLES OF INTERPRETING SKIN CONDUCTANCE RESPONSES

3.7.1 Mary Listening to *Romeo and Juliette*

Mary has a large skin conductance response at the start of the main theme of *Romeo and Juliette*. This response is much larger than any other time during the concert, so the affect was not likely due to a subtle shift in posture, which she likely did throughout the concert. When I interviewed Mary, she put her hands to her heart and said she loved that part of the music. I note this as a moment that the music excited Mary. When I asked Mary why she loved *Romeo and Juliette*, she told me it was a song she knew, so I note this as a case where familiarity played a role in Mary's reaction.

3.7.2 Mason Talks to his Mother about Technic

While building LEGO technic Mason says, “I’m building LEGO Technic by myself.” Correspondingly, we see a peak. I first ask myself whether this response is due to physical movement or breathing. He does start talking, but 7 seconds beforehand he was also talking to his mother with a much smaller skin conductance response at that time. This suggests that what he said to his mother had importance in itself. Later on from my observations and interview with Mason and his mother, I learn that he likes to build with his mother. Mason was likely anxious about building Technic. However, it might be that he was worried that his mother will not help him as well. Alternatively, he might be excited at the idea of building Technic too, though this is much less likely as he quits soon after this announcement. I label this response with all of these possibilities, and from a design perspective conclude that Technic may be perceived as difficult and overly-challenging. I also note that children may need new ways to invite their parents to play with them, a topic that I hope to explore further.

3.7.3 Buying a Blender with Multiple Stimuli

This example is a hypothetical example, pulling from my retail work, but it covers many difficult-to-interpret situations. A person is shopping for a blender and an employee talks to him about the blender warranty. Simultaneously, the customer physically moves his head and corresponding attention to a new blender. A Black Eyed Pea song’s loud bass section starts thumping over the public speakers during this time. Concurrently, the individual’s cellphone goes off and starts vibrating. This type of scenario is difficult to interpret correctly. I know from past research that cell phones going off during shopping often create a large response. Even with the cell phone, I am not sure if the shopper is annoyed or embarrassed by the ring or just orienting to the vibration. I am sure all of the other factors could also play a role as well, but I cannot conclude cause with high specificity, as the response could be due to the cellphone alone. If the shopper reported they were frustrated that the employee was talking about insurance, I would note this as important, but I would not conclude the skin conductance was due to this frustration.

3.8 SYNTHESIS

After the skin conductance responses are labeled or discarded, I conduct synthesis. While synthesis of an interview can be difficult (so what were the main points we heard from him again?) synthesis with skin conductance is straightforward. I take all the labeled skin conductance responses and cluster them together based on what occurred. I can then group these similar responses into larger clusters. With the Blue Man Group data, I noticed that people tended to react whenever an audience member was embarrassed, so I grouped the “video on person coming late,” “audience member on stage eating Twinkies,” and “participant catching a marshmallow” all together as the general idea of audience participation arousing participants.

During this time, I recommend mixing the qualitative findings with the skin conductance findings. During the New World Symphony, novice participants reported that they loved seeing the visuals despite no skin conductance responses occurring at these moments (the visuals were always on display, allowing habituation to their presence). I clumped this with “responses occur when novice members focus on players” and “responses occur when audience members whisper to one another” in the larger category of “the music is not enough sensory input for novice listeners, they want something else to focus on.”

3.9 EXPLORATION AND PROTOTYPING

Of course, pinpointing the causes of stress or frustration does not directly lead to a better design. The second step is to brainstorm possible, new solutions to these problems. Sensors can be used in this setting as well. When I identified that typical classical concerts were boring after five minutes, I started looking at my own skin conductance during rock concerts and other people’s skin conductance during the Blue Man Group. As a small example, I noticed that when the Bluemen acted in the audience rather than the stage, the audience members tended to have a strong skin conductance response. I listed this technique as one way the New World Symphony could help alleviate the boring parts. For the LEGO tablet project, I measured children’s skin conductance while they played *Minecraft* and other app games (their skin conductance was much higher). This observation helped show that building LEGO sets plays a different emotional role than *Minecraft*, and when designing the tablet experience, we should be sure to keep with the calming, meditative experience that building LEGO sets normally brings. However, if we want to excite children before or after they build, we could copy the social element of tablet games, which tend to create skin conductance responses.

Once a few possible solutions are decided, psychophysiology can also help with the prototyping stage. Seldom does a solution to change an emotional experience work on the first try. When I work on prototyping for the retail space, I create a work process of a half-day building and a half-day observing and measuring shoppers in the retail space. We repeated this process for two weeks: noting the emotional experience shoppers have in the afternoon, and then changing the setting in the morning to address the new found needs. This setting makes analysis more difficult, as time is crunched. I recommend having two researchers during prototyping – one person to conduct the interviews, another to annotate the skin conductance live. Additionally, people will react to various features that are only half finished. For example, “I can’t find the blender model number!” will result in “noisy” graphs. From an interpretation standpoint, this is not a problem; researchers can ignore those responses that are outside of the prototype question. However, communicating that some of the emotional responses should be “ignored” to clients may be difficult. If evidence is important during prototyping, I suggest having 3 or 4 days at the end to measure people’s skin conductance in a finalized environment, so enough responses can be collected around the revised idea. If done correctly, this prototype data can be extremely powerful, as researchers can show the before and after effects of their interventions.

3.10 COMMUNICATION

3.10.1 Introduction

As I wrote the Communication Section, at times, I felt I could have titled it “How to Lie with Graphs” as much of this section is about altering the graphs to communicate appropriate messages. However, in reality, interpretation of skin conductance is complex but still needs to be accessible for those who make decisions. Data should be shown in a way that helps users see what the expert is seeing. Too often “the expert” removes people from the data and evidence and gives them a “trust me proof” of $p<0.05$ or a bar graph simply stating skin conductance proved Product A was more stressful. Much like an ethnographer chooses key quotes, data should be presented in a carefully constructed narrative, which helps people to understand from where an interpretation comes. In the future, other complex methods of data science will be fraught with finding ways to explain the complex processes in a way that every day decision makers can understand.

Perhaps a simpler theme is that analysis and communication are different. When one has a critical eye and patience, the data do not need to be formatted in a clear, suggestive way. In fact, it is dangerous to do so. However, when non-experts view and interpret the data, the graph’s purpose changes, and graphs should help communicate a clear, accurate story that will not be misinterpreted.

3.10.2 Explaining Skin Conductance

After six years, I have explained skin conductance to thousands of people. Most often I have 30 seconds to explain a phenomena that Boucsein has written a 624 page book on (2011). This puts me in quite a pickle: How do I explain this rather complicated phenomenon in a way that is simple and accessible and allows people to understand what we do with psychophysiology and design. My go to message is:

"When a person becomes stressed, frustrated, or excited, the technical term here is 'physiologically aroused,' and then we will usually see skin conductance responses like this."

This sentence emphasizes the tying of emotional response to skin conductance and how the response can occur for a variety of reasons. However, listeners often fall into believing that a skin conductance response means that a person is having an emotional response, which is not true. I then follow with,

"When doing these studies in the real world, we have to figure out which possible mental state created a response. To do this we use observational and interview methods to help narrow in on what occurred."

I then give a clear example. I like talking about my skin conductance when using a self-injection device, as it has clear skin conductance responses, and no one questions whether a shot is stressful or not. Not to mention, my fear of shots is pretty entertaining.

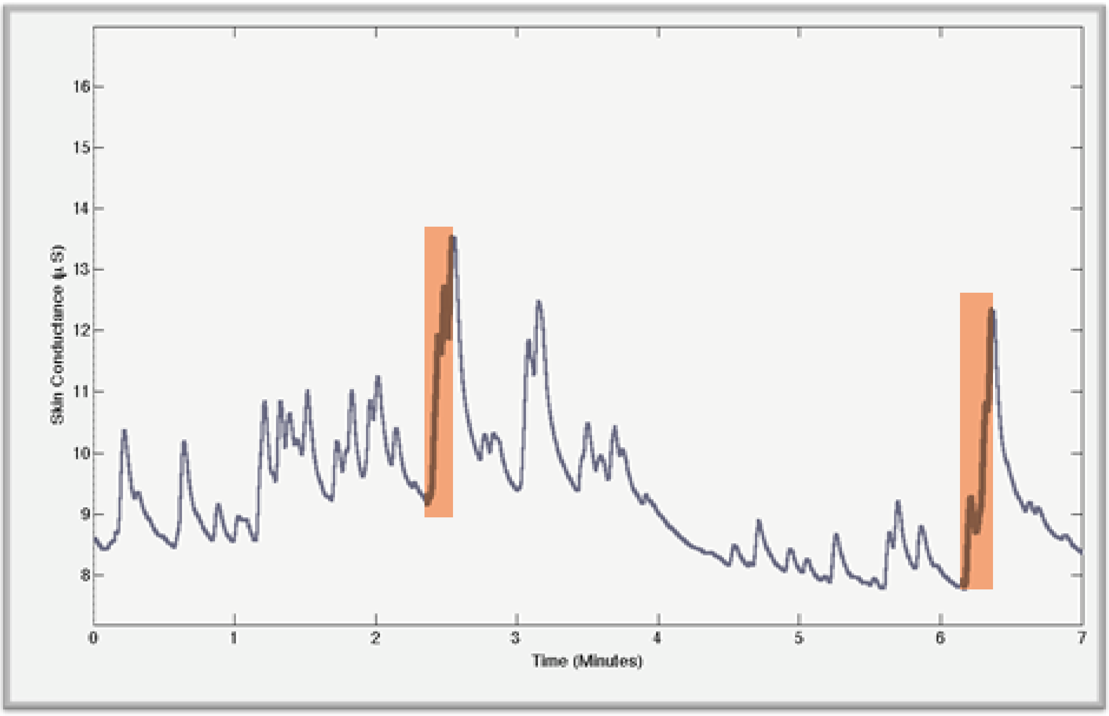


Figure 3.18 My own skin conductance as I learn about a self-injection device. This is my typical introductory to skin conductance example.

Unfortunately, listeners judge the conclusions from the research immediately, rather than the complexity of the interpretation. It is my belief that the first example shown to business people should not be directly relevant to the clients, and especially should not show data about the client's products or services first. Showing result data will make people defensive and motivate them to start looking for flaws in the method, before they have even seen it work.

When speaking to non-scientific audiences, I do not use the word arousal and stick with the emotion words they know – stressed, frustrated, excited, and overwhelmed. I emphasize during my talks that these reactions come from long interviews and observations in addition to skin conductance.

3.10.3 Graphs

The Skin Conductance Level Obsession Problem

Since we first learned about graphs as elementary students, we have been trained to look at level rather than slope of graphs. This habit creates problems, every time I present a skin conductance graph. Most people default by looking at skin conductance level. This number is relatively meaningless from a design perspective. A person's skin conductance level of 13 uS is not comparable to another person's level of 2 uS. People have different amounts of sweat across time and places as well. The first challenge of graphing the data is to move people away from this focus on level.

I shift people away from focusing on level in two ways. First I deemphasize the absolute skin conductance level. I make the text tiny and unnoticeable. Of course, making it too tiny will frustrate viewers as they fiercely try to read cryptic numbers. Additionally, researchers will want to see the numbers to make sure the range is trustworthy.

The second way I emphasize slope is by highlighting upward slopes as important. This highlight helps people focus on amplitude, as the color bar has a clear start and stop. Additionally, this highlight method removes people's attention from the downward slope. Skin conductance is more like Morse code, sending responses every once and a while, creating the upward responses. The slope down is an exponential decrease, like a capacitor. This is a consequence to the elevated skin conductance level, and does not provide additional information for analysis. Some researchers suggest area under the curve may be a strong marker of arousal as well (Ben-Shakhar, 1985). This method would include the downward curve. But as many ambulatory responses do not return to baseline during ambulatory sessions, this method is still impractical.

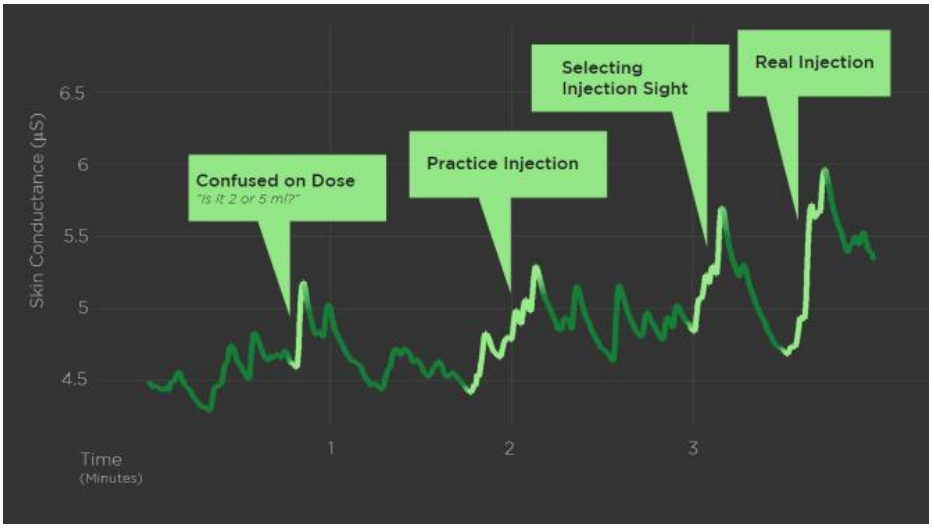


Figure 3.19 Skin conductance responses are highlighted to emphasize the slope nature of the signal.

Y Scale

When clients or other researchers send me data, it is scary how often it looks like the underwater level from Mario.

Figure 3.20 Underwater level from Mario.

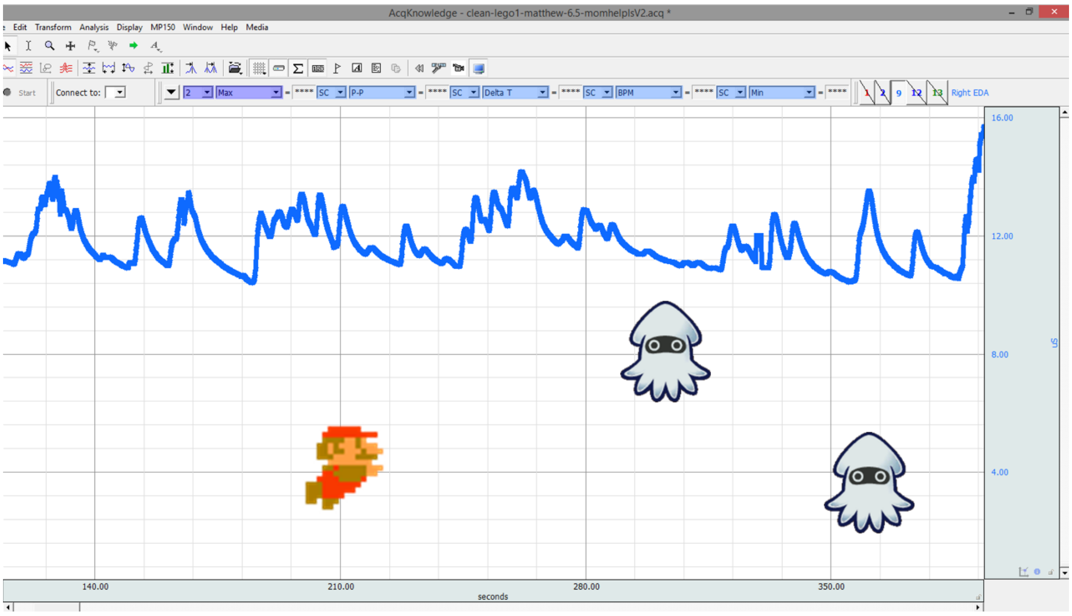


Figure 3.21 Skin Conductance Graph with Y minimum set at 0. Mario and Bloopers added for effect.

This graph minimizes the responses and emphasizes the skin conductance level. The level of 0 uS is rather meaningless, and Y minimum should not be set to 0 without reason.

Altering the Y scale has drastic effects on viewers, and will tell different stories depending on how it is manipulated. As people visually more than numerically compare charted data, the chosen scale determines the interpretation. Take for example Mason announcing that he is building alone, the moment that follows, and then Mason getting up and walking to his mother. All three of these events are meaningful, yet the Y scale can make them appear significant or insignificant.

Figure 3.22 A child talking to his mother about "being alone" appears stressful.

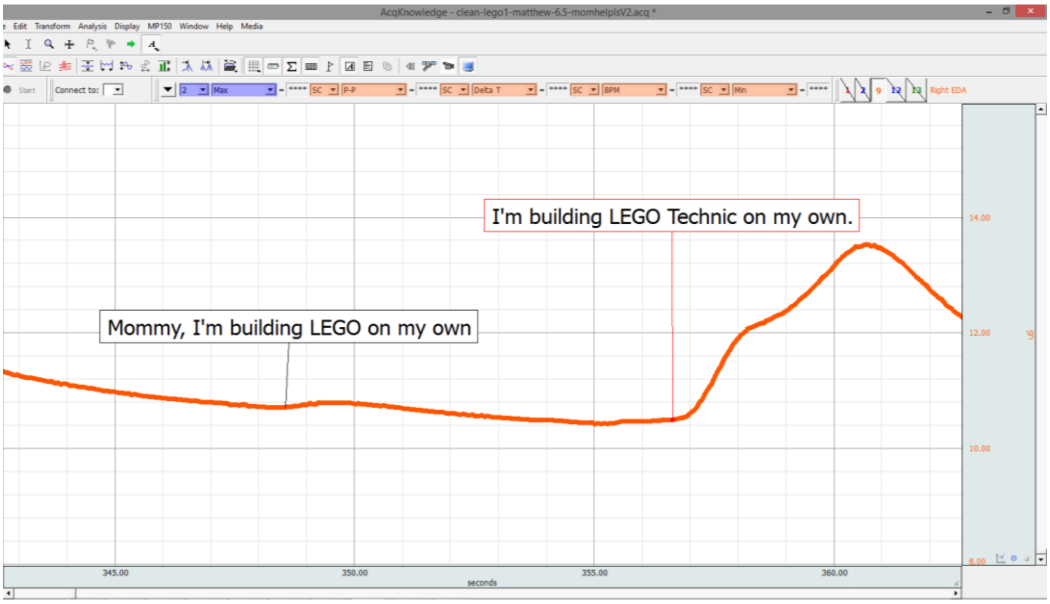


Figure 3.23 A child talking to his mother about "being alone" does not appear stressful, but talking about building Technic appears stressful.

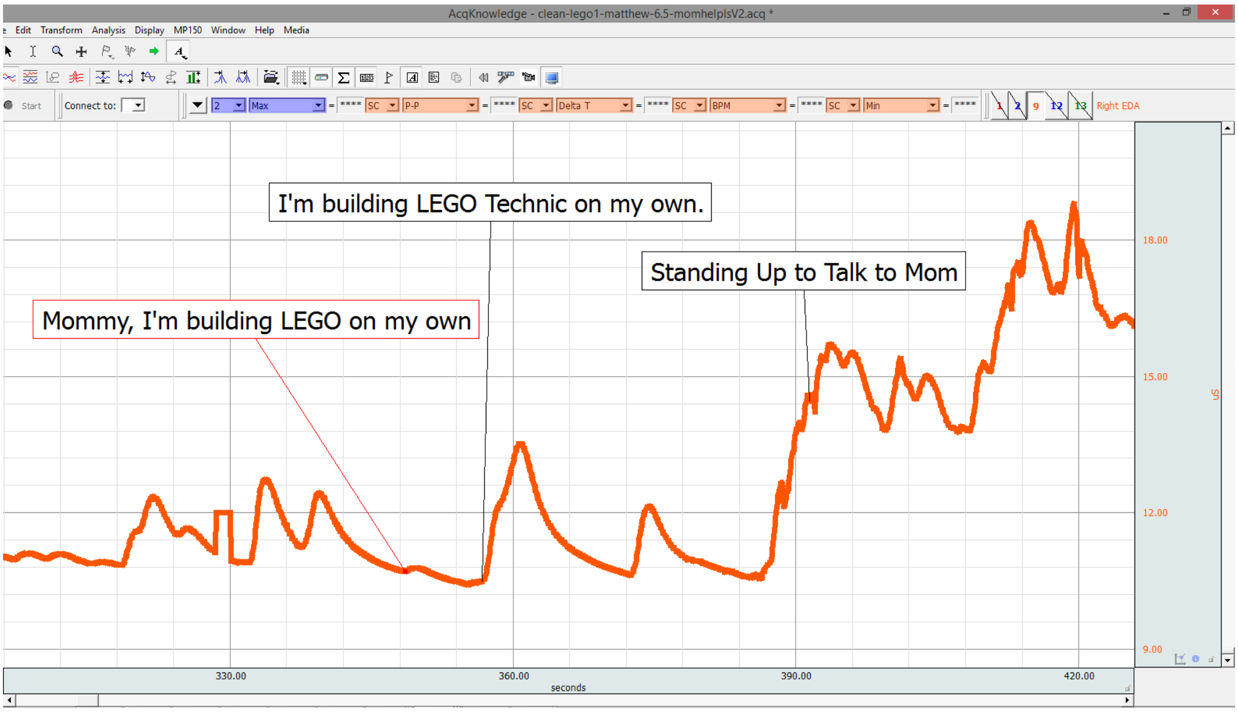


Figure 3.24 Neither talking about Technic nor building alone appears stressful compared to the increase of skin conductance to standing up.

From a research standpoint, I know both times Mason talks are important as the response increases by more than 0.05 uS. However, people will not consider all responses significant but rather look at visual size comparison. When determining which scale to use, I ask what is the message I want to communicate. In this case, I wanted the LEGO group to understand how hard it was for Mason to ask for help. We can see him a little worried with the first response, and by the second response he seems to be panicking. In this way, I set the scale so both responses could be seen, but the second one was clearly larger than the first.

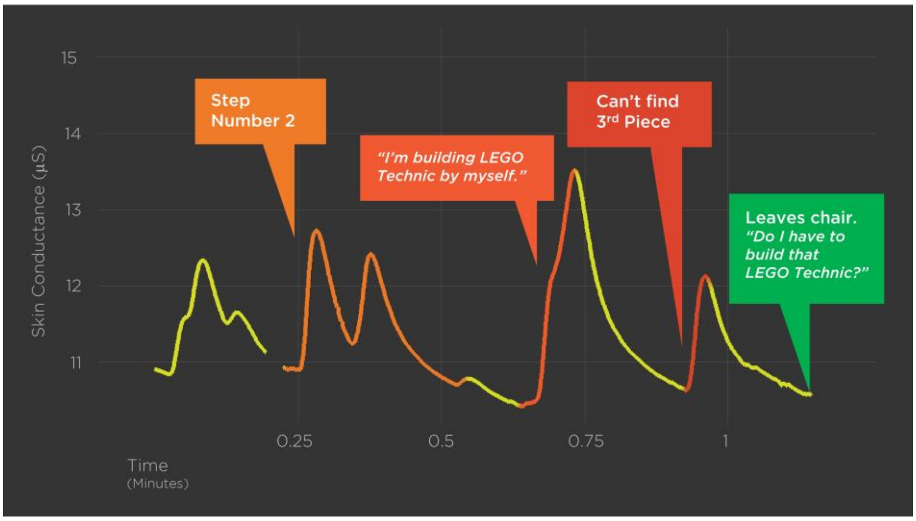


Figure 3.25 Finalized graph, emphasizing Mason talking about Technic.

The Changing Baseline Problem

Unfortunately, even with highlighting responses, people will still evaluate by the skin conductance level over amplitude. This becomes especially problematic if skin conductance level changes across time.

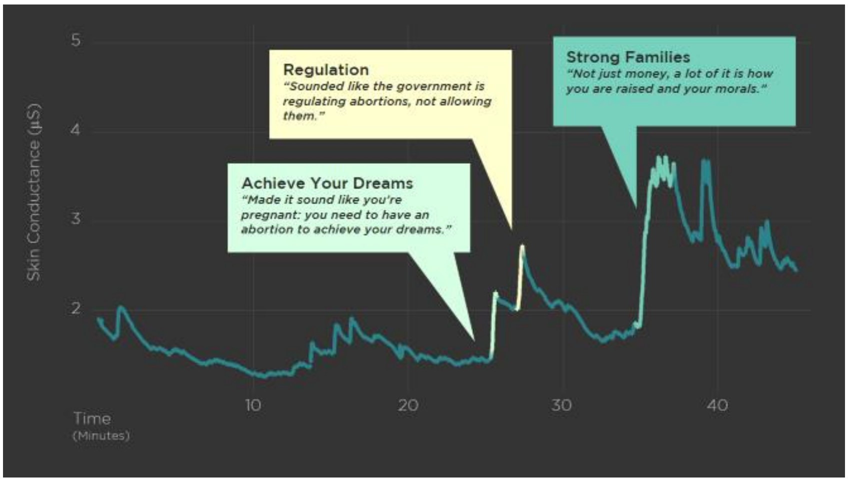


Figure 3.26 The response to "Regulation" is higher than the response to "Achieve Your dreams," but they have similar amplitudes.

In this example, a woman's response to "Achieving Your Dreams" and "Regulation" have similar amplitudes. However, the second one has an increased level, so people will incorrectly interpret it as twice as strong. After all, it is higher up on the graph. However both responses are about the same in terms of amplitude. At IDEO, I was asked to translate these into bar graphs, so they can be more easily compared.

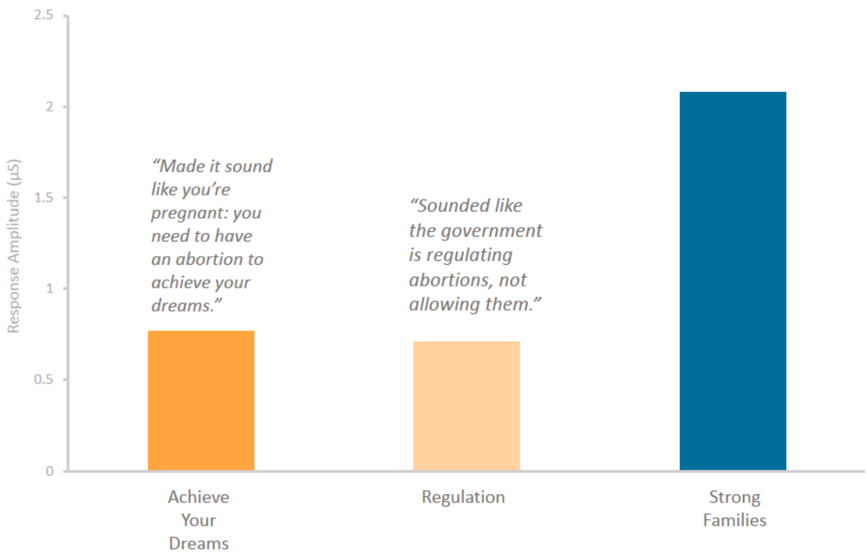


Figure 3.27 Bar graphs remove the reader from the skin conductance.

While this bar graph does clearly emphasize the amplitudes, this data becomes too abstract. Skin conductance allows people to more directly observe an individual's emotional experience. The biological signal creates a sense of intimacy. If the skin conductance is translated to a bar graph, one could just as well say an effect occurred without showing any graph. In bar graphs, there is little visual difference between skin conductance data and data from a survey asking people how stressed they were at a given moment. One possible remedy is to show both the skin conductance and bar graphs, but this over complicates the description, and people tend to only want the minimal amount of information necessary.

The Artifact Problem

Movement can create large response artifacts as well. If the sensors become temporarily disconnected from the fingers, a quick jump may be present. These are easy to avoid during analysis, but are an eye sore during presentations.

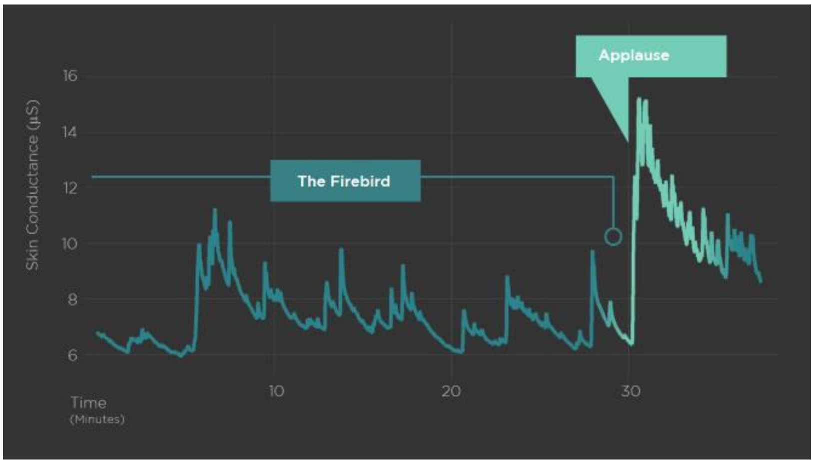


Figure 3.28 An expert listener reacts to a classical concert.

In this graph an individual reacts during the classical concert of *The Firebird*. The applause afterwards dwarfs these smaller responses visually. However, this last response could have occurred as the man stood up or clapped, making its "larger" nature not necessarily accurate in interpretation.

Even if people are told to ignore the large hump, that hump will be visually daunting, and suggest the other response effects are less meaningful. If I ask how tall a participant is, if he is standing next to a giant, he is going to look smaller, even if I tell the viewer to ignore the giant.

Figure 3.29 Even Shaquille O’Neal is puny next to mighty Cthulhu.

If the skin conductance shortly returns to near base line levels, I will remove the artifact from the graph, having a small gap. But often times, level does not return, and leaves an unexplained shift in the graph. And when skin conductance reappears, it will be at a higher level, likely having the viewer assume the individual is highly aroused, which may not be the case.

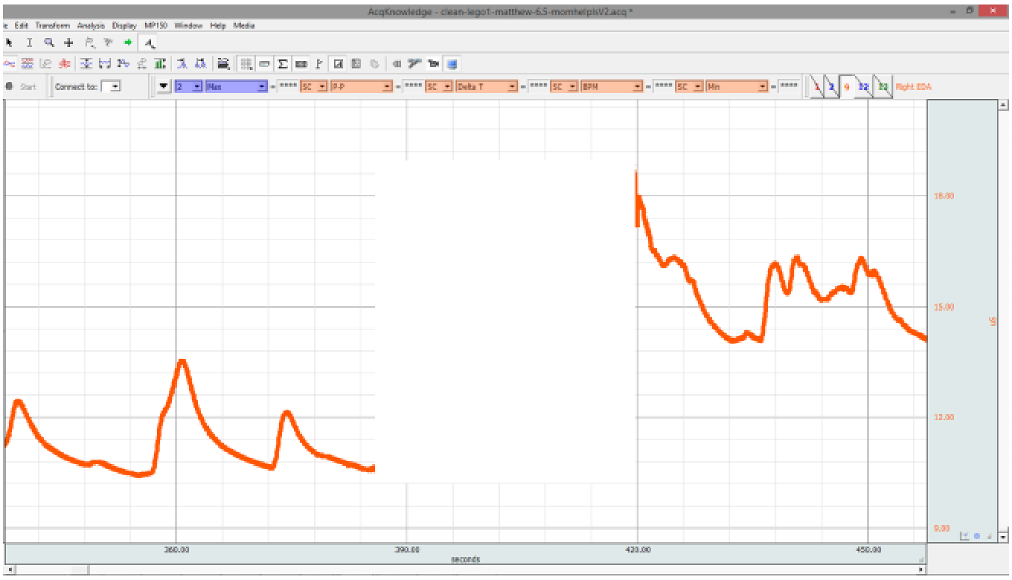


Figure 3.30 Removing movement artifacts can still create problems, as the skin conductance level is now mysteriously elevated.

Time Scale

Like the Y Scale, the time scale will alter how people interpret the data as well. Much like a full Y scale, too often I see plots of the entire time.

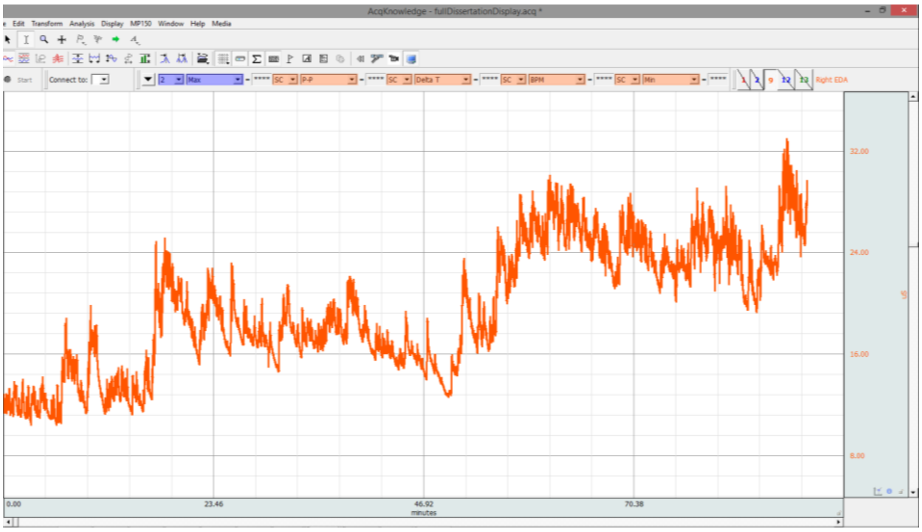


Figure 3.31 A large time scale provides too much information simultaneously.

Showing the entire session creates two challenges. First skin conductance responses are so tightly packed, that it becomes difficult to see them, and if a person does try to read the responses, they will be overwhelmed. Secondly, baseline will shift during a session as people react more or less. This shifting baseline could diminish single responses, making many of the responses seem tiny. With long time frames, people will likely focus on changes in level. Too short of a time period also creates a problem.

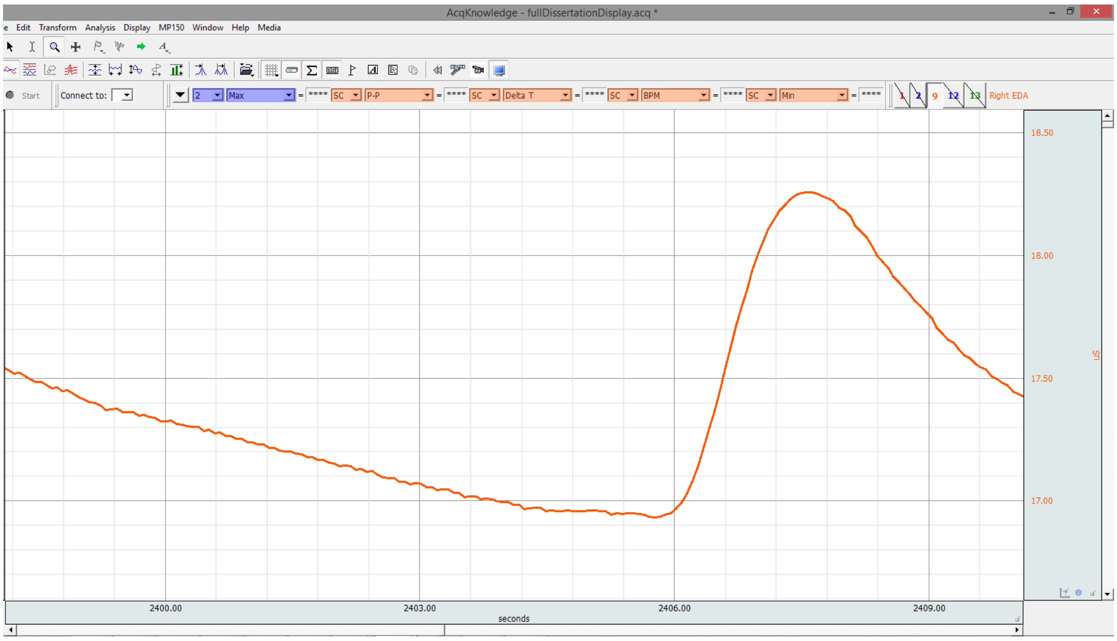


Figure 3.32 An overly zoomed in time scale removes the response from context. Can you figure out where this response is from the previous graph?

In this plot all one can see is a skin conductance response. A response is rather meaningless without something to compare it to. Is this supposed to be bigger than another part? Did this occur surprisingly after a long period of no responses? Graphs with only one response can be misleading and uninformative. In general, I create graphs that show a clear change occurring across time. Ideally, I graph multiple responses to help tell a longer narrative of multiple events.

Business clients sometimes ask to see the skin conductance response across a longer period of time. This makes sense in part – are the responses I show negligible to responses earlier or later on? In ambulatory settings, different size responses are likely to occur. Imagine a customer looking at blenders, walking to a checkout, and then checking out. There is no reason to believe that the responses between the events should be the same nor does it make smaller responses in browsing less important than the larger responses while walking.

To help show the larger picture, I used to put a small graph on the bottom of the response over the entire time, with the small part highlighted. This graph never seemed to be viewed and added a lot more information – two different time and Y scales and multiple responses distract the viewer. No one seemed to comment on these graphs, but did question them, so I removed them later on.

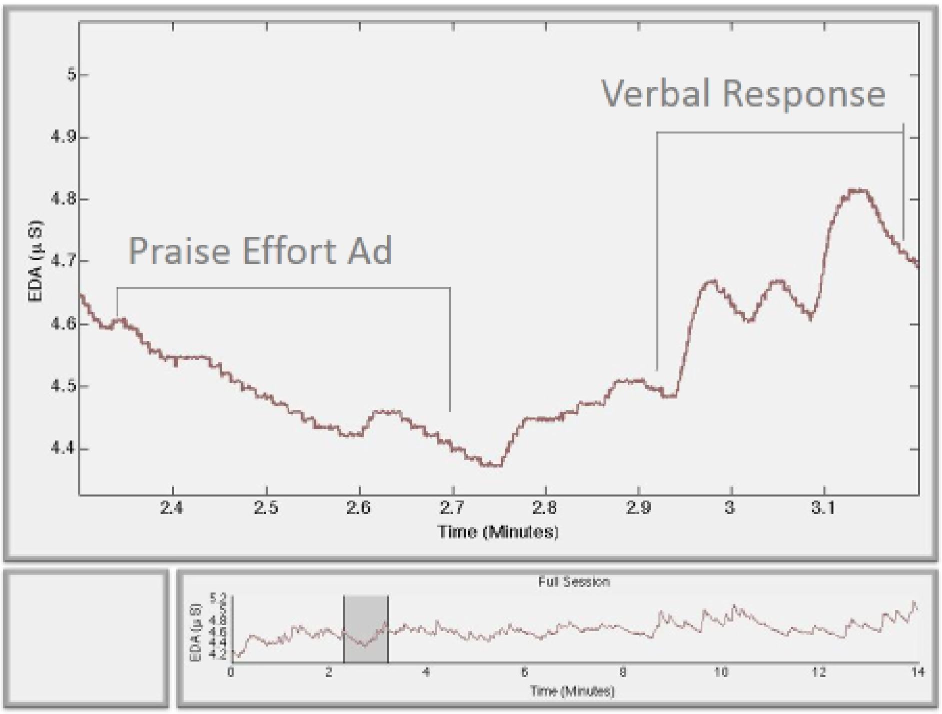


Figure 3.33 An unfortunate attempt to show zoomed in data and larger responses simultaneously.

Annotations

Skin conductance without described context is difficult to interpret.

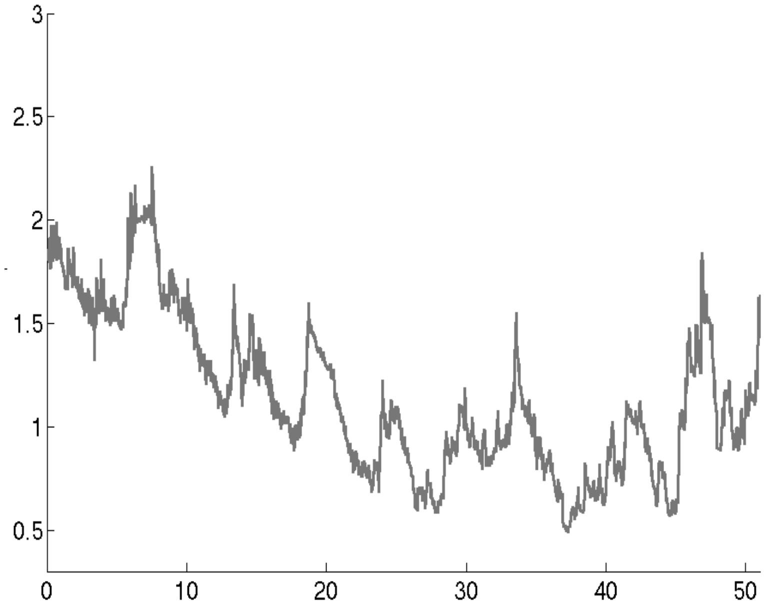


Figure 3.34 According to this graph skin conductance changed, but why is unclear.

Viewers want to know why a skin conductance response occurred. We need to visually communicate this story in the graph. One approach would be to label an event above with lines marking the onsets.

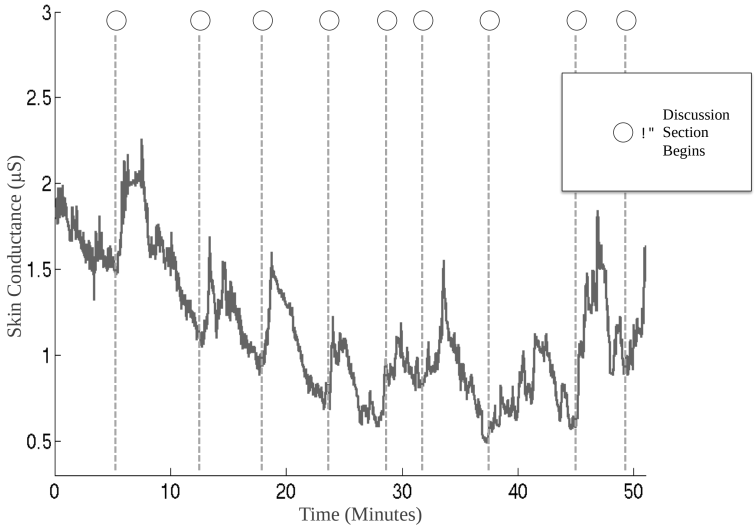


Figure 3.35 Skin Conductance with discrete markers depicting when events occurred.

The main problem I have with this method is that stimuli are seldom constant across an ambulatory setting; each time a shopper looks at a blender the experience is likely a little different, so event labels are misleading. And even if the stimuli are relatively constant they are seldom small events in time. Usually events occur over a larger period of time.

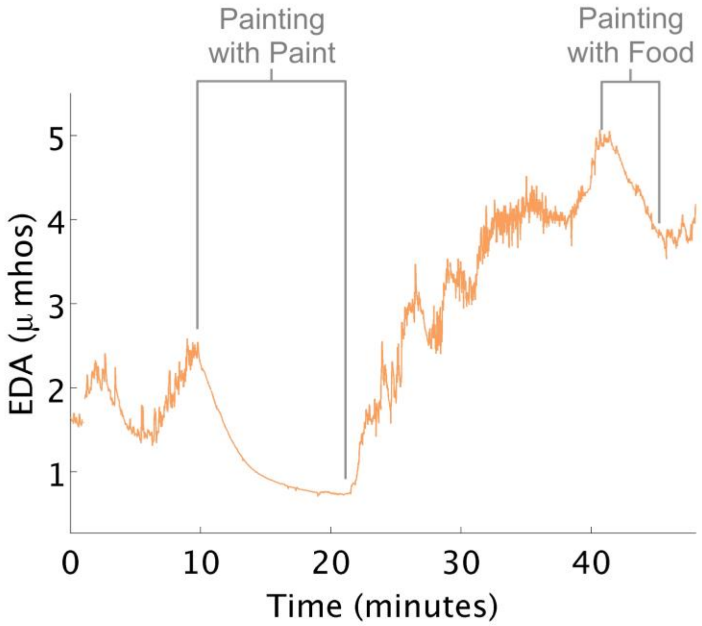


Figure 3.36 Showing when events occurred over a longer period of time.

One nice part about labeling events across time is researchers can determine how much time occurred before or after a response. I find this more interesting from a research perspective than a communication perspective. People I present to tend to not be interested in the exact specifics of when events start or stop, but more why a skin conductance response occurred. With longer time frames, responses begin to be confused as well. For example, when Anna is in the ball pit she stands up half way through, so the viewer will become confused by the ball pit being labeled as having both an up and a down response.

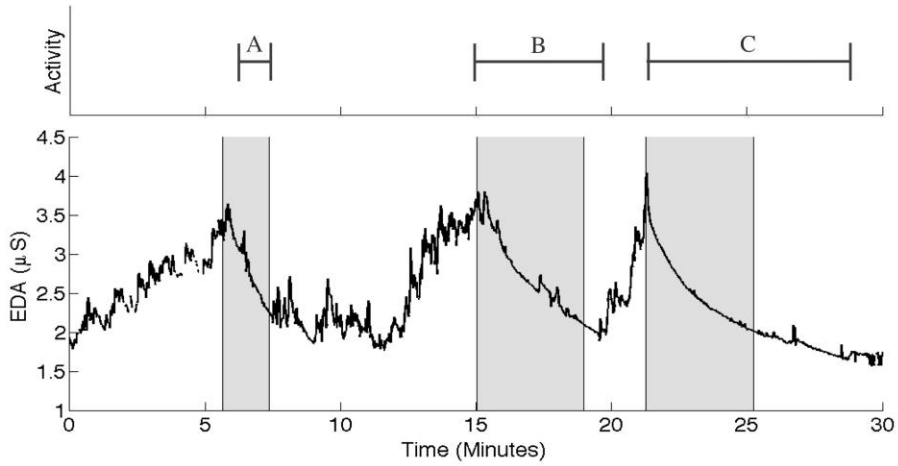


Figure 3.37 The skin conductance of a child lying in a ball pit (Section A and B).
She was in the ball pit longer than the A section, but began moving around and leaning over the enclosure, resulting in her skin conductance increasing. If I were to widen section A to all of the time the child was in the ball pit, readers might be confused why the response started changing (Hedman et al., 2012).

One case where this highlighting of event times works well is when showing NS-SCR frequency. Being able to show a multitude of responses appears nicely with highlighting.

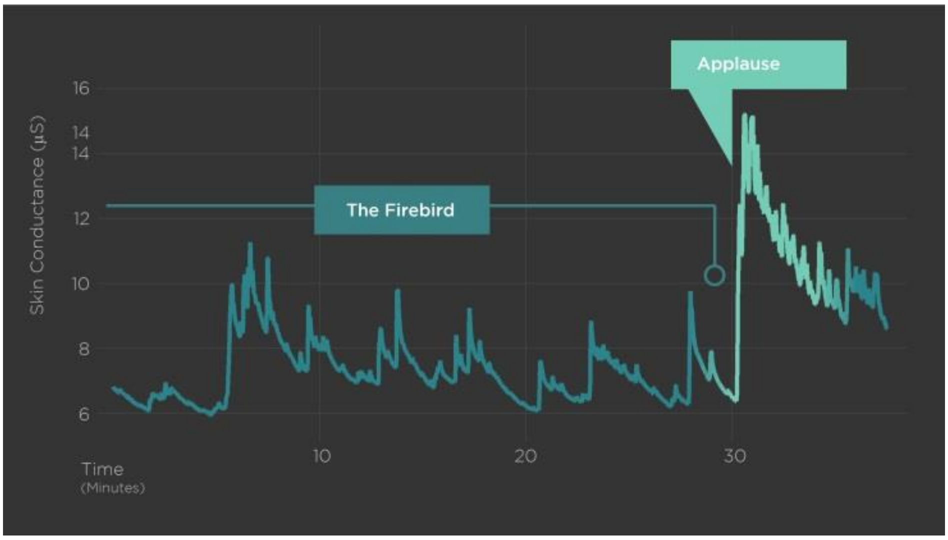


Figure 3.38 By highlighting the entire time *The Firebird* played, I was able to show how the audience member responded throughout the music (see Chapter 5).

With advice from IDEO, I started labeling the responses themselves rather than the events.

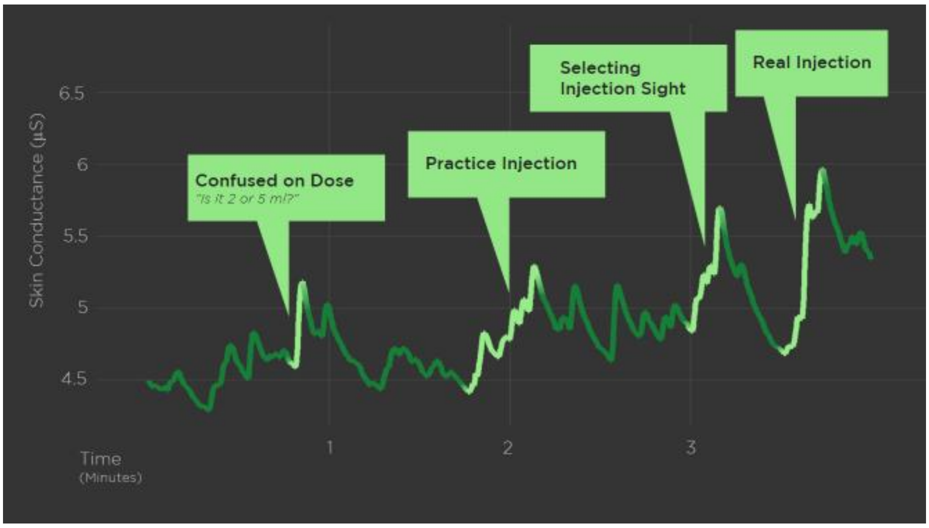


Figure 3.39 Annotating Responses.

This helps link the stories to the skin conductance and highlights the slope, what people should be looking at. While this graph quickly communicates the findings, it does require trust from the reader. It is unclear how long the stimuli occurred for or when it started and stopped. Generally these issues are more interesting for the science part of analysis. This label strategy also allows for easy addition of quotes as well. These quotes help bring the participant to life and explain how we determined the interpretation behind a response. Including quotes also emphasizes the qualitative nature of these interpretations.

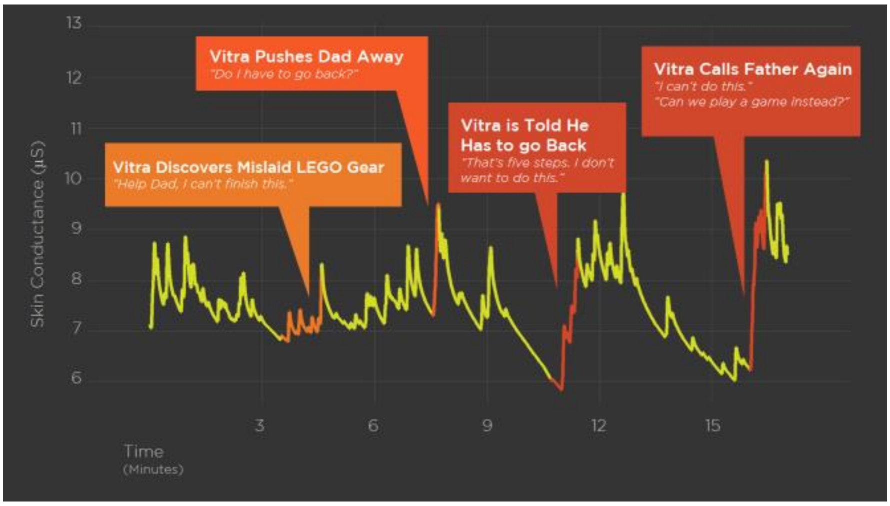


Figure 3.40 Adding quotes to annotations.

3.10.4 Video

When I started showing therapists graphs of children's skin conductance, a peculiar request popped up: the therapists wanted to see video with the graphs. After all, there was so much going on during a response. Noting Mary was stressed as she approached the swing was not enough. The therapists wanted to hear the anxiety in her voice, to see her struggling to get on the swing, to observe what the therapist was trying to do, to be immersed in the context. Videos can be powerful ways for viewers to dive into the setting around a skin conductance response. In fact, during my hour-long presentations, I spend a good 20 minutes showing videos and discussing people's reactions. Seeing skin conductance change with events occurring helps solidify the data in reality and helps viewers ground themselves in the context.

Of course, I discovered all of this long after my initial tests in the occupational therapy center. The original videos I made looked like this.

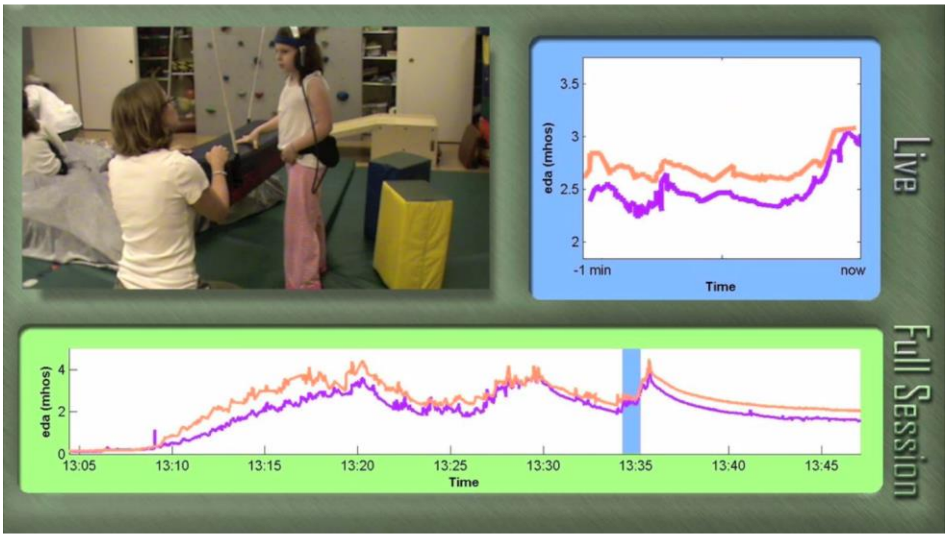


Figure 3.41 Original attempt to combine skin conductance and video.
YouTube address: http://www.youtube.com/watch?v=epKEKmxvWFg

In the above graph, the child's real-time skin conductance is graphed in two windows. The top blue window shows the last minute of skin conductance and the bottom graph shows the child's skin conductance across time.

The latest videos I now make look like this

Figure 3.42 Latest attempt at combining skin conductance and video.
YouTube Address: http://www.youtube.com/watch?v=tYUKpXg-C8g

In this video, skin conductance is overlaid across video. Only one response is shown and skin conductance responses are highlighted. A ball highlights the child's current skin conductance level, and quickly rises up with responses.

Y Scale for video

Similar to static graphs, moving graphs should emphasize the current skin conductance response amplitude, and the scale can emphasize or hide that amplitude. Find a Y scale that helps emphasize the responses as much as possible, and stay with that scale across time. This may be difficult as baseline can shift over time, and the shift in average skin conductance level could make skin conductance responses seem minimal. In these cases, a dynamic Y range that shifts its center and range with the current content would work better, which would deemphasize the shifting baseline. Alternatively multiple short videos could be created, all with set Y Scales.

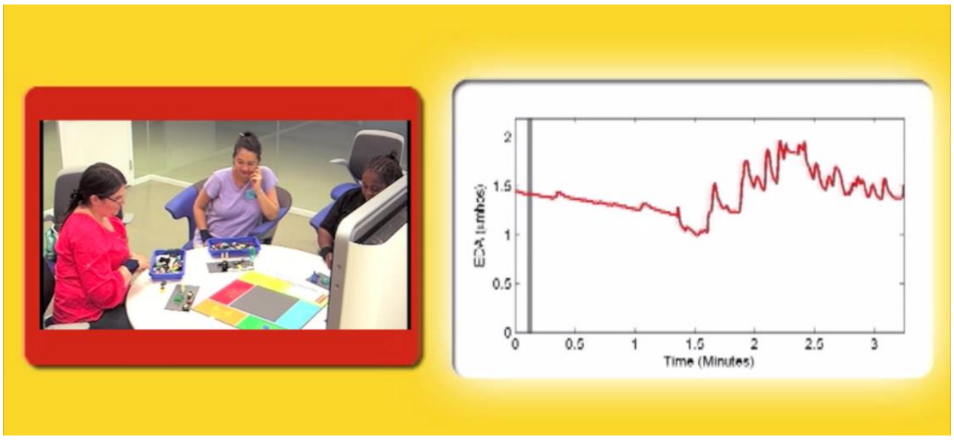


Figure 3.43 If Y Scale is stagnant, some responses will be viewed as small.

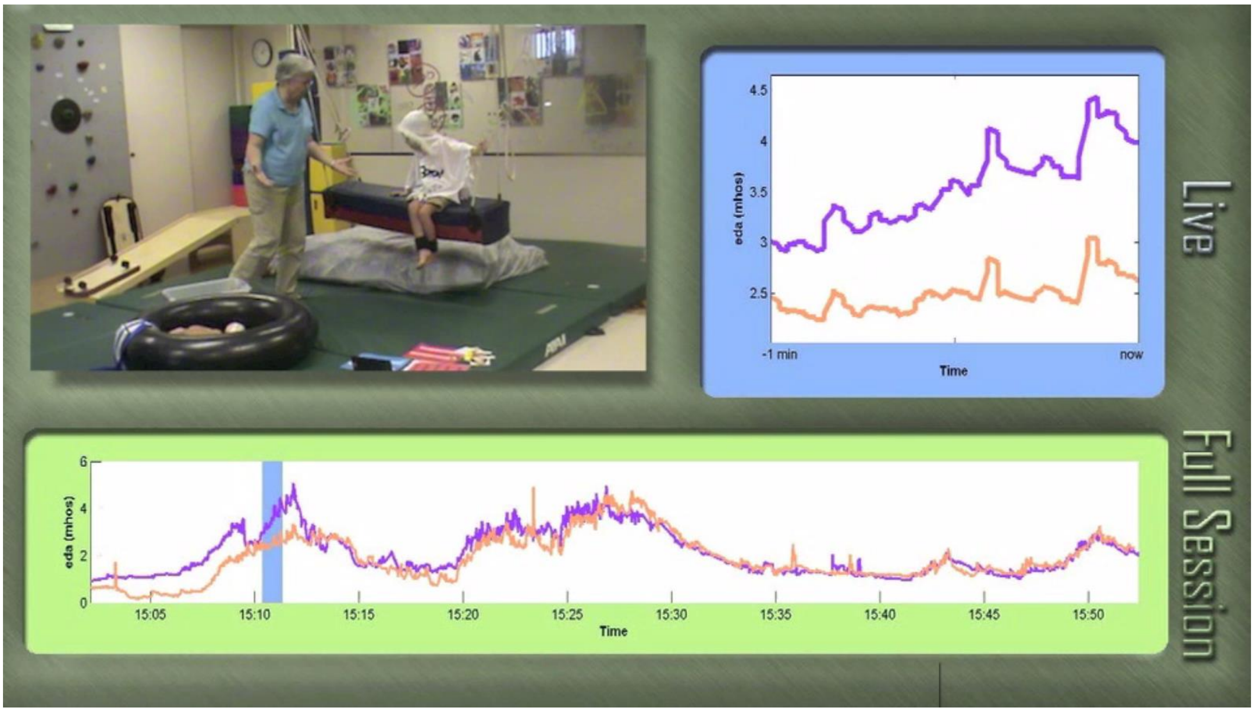


Figure 3.44 At this point in time, the child's skin conductance is increasing considerably, but the dynamic scale down plays that increase.

One trick I have enjoyed is extending the window for local maximum and minimum by a minute on both ends. This keeps the scale from quickly shifting back and forth and also helps large changes come in as large changes, rather than a similar-looking line with a changing scale.

Dynamic range also has the problem of emphasizing differences that are small, specifically if the person's skin conductance is flat with no responses. During these times, I make sure the difference between minimum and maximum skin conductance remains at least 1 uS.

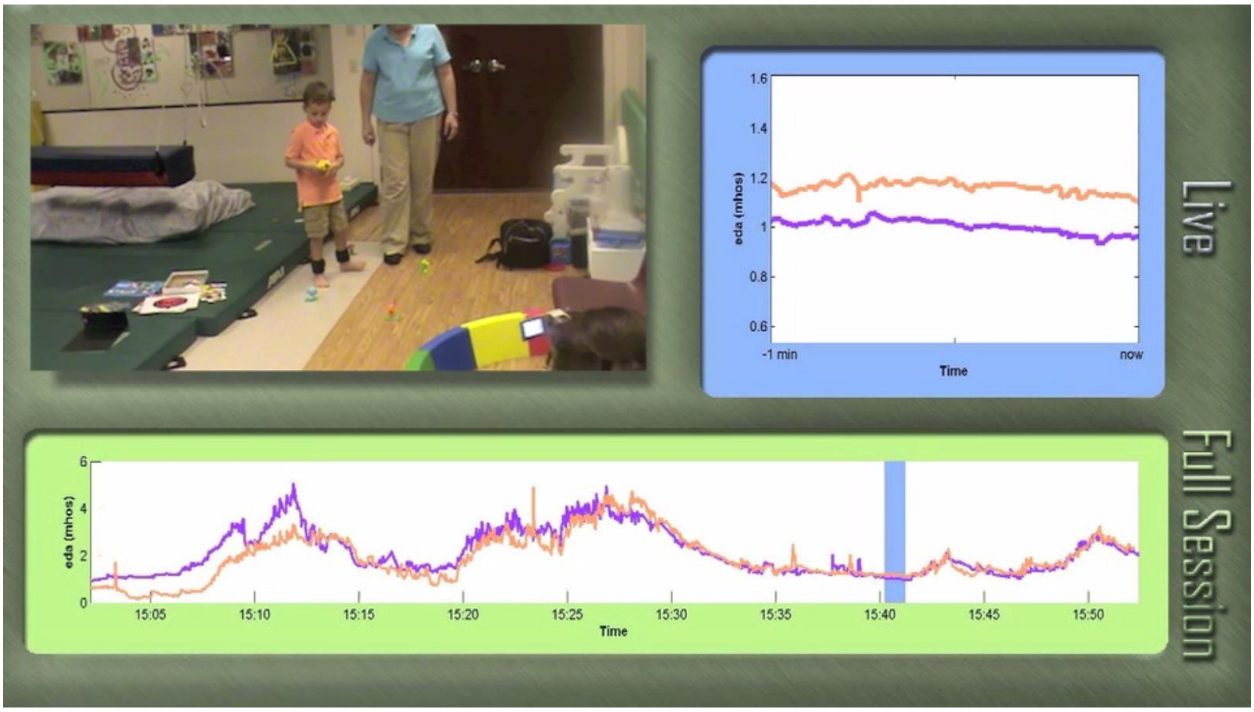


Figure 3.45 Set a minimum range for the Y axis to prevent noise from a flat signal being emphasized

Time Scale for video

I want the viewer to only look at one response at a time, but I need to give the graph enough space for the viewer to see how the latest response differs from the previous ones. To do this, I keep 10 seconds of data present, with the right end being the current moment. I put a circle around the end to emphasize this current moment. This circle has the added benefit of bouncing up every time a response occurs, which is a strong visual signal. A benefit of having the graph data react in real time is that time labels are unneeded – the movie time is the time label. Sometimes, I shift the video by 1 to 2 seconds, so people do not have to mentally calculate the small physiological delay in skin conductance response. This becomes more important when the graph is overlaid with the video and people can watch the graph and video simultaneously.

Previously, I put a longer graph below to help locate where the response is over time. Like the additional static graphs, this adds a large amount of information for first time viewers, which has confused them in the past. This method does help in locating key moments as a researcher though.

Another option is to have a line that moves across a stationary graph as a way to indicate time. This was my standard approach for a while as well. This method is great for analysis, as the researcher can view responses over time and zoom into areas of interest.

However, by having a line move across space, the viewer has to switch from viewing events in time (video) and spatially (graph).

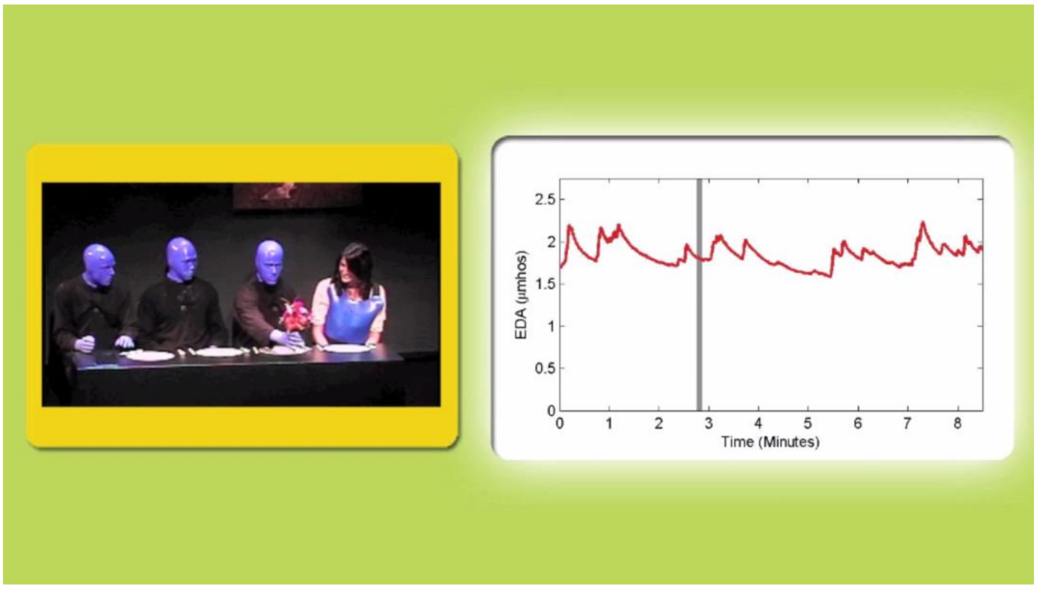


Figure 3.46 A bar moving across time asks the viewer to translate the play time spatially, which may be distracting.

Video Integration

In my initial videos, one could see how obsessed I was with the graphs – the video was a little part of the display. Over time, I have come to realize that video is a more complex, detailed feature for viewers, so it needs to be emphasized. I recommend making video take up as much of the screen as possible. In my current iteration, it takes up the entire screen.

I also recommend that the graph overlay on top of the video, so people can view response and video simultaneously. Otherwise, viewers will be forced to switch between skin conductance and video, not allowing them to fully focus on either. One challenge to this approach is the line may be covering important information, like people's facial expressions, so researchers should take care to place the signal in a non-obstructive location. Transparent skin conductance lines also help. When overlaying the videos, I like to make the current point (the ball) line up with the participant's face. This is the space most people will focus on, and makes it easier for viewers to stay focused on both pieces of data.

Video Discussions

When I was presenting at the Design Research Conference, I was asked to give an interactive presentation. I thought it would be fun to have researchers analyze the data as I do.

I posted a video onto YouTube and had everyone visit the site. This method broke when I realized 200 people would not be able to access YouTube simultaneously in a conference room. So I improvised and showed audience members skin conductance of a mother as she taught board games. I showed 30 to 60 seconds of a key part in the interaction and then had the audience discuss what they think they saw with each other, and then with the whole audience. I then repeated this process and showed another clip and had them discuss again. When people shared their thoughts with the general audience, I would confirm some of their ideas with other examples, observations, or quotes from the study.

While a complete improvisation, that presentation was well received. A Core 77 article highlighted my talk, asking for me to send along the video we analyzed. Clients who wanted to work with me asked for the video to show their bosses. I believe this interactive method works because it creates a sense of ownership. People are not passively being told what to believe, but they are asked to think hard about what is going on during the videos. Viewers then create their own meaning and thoughts and share it with one another. This method helps people own the data. An additional bonus is that the presentation emphasizes the qualitative and interpretive aspects of the process. People also do not ask me to justify my findings, because they were able to see the participants get stressed or frustrated, right along with the video.

3.11 CONCLUSION

If you are a psychophysiologist, this chapter may have made you uncomfortable. I have shifted away from most of psychophysiology's traditional laboratory methods. I understand that there is a need for robustness in scientific conclusions, which thick psychophysiology does not provide alone. However, as psychophysiologists start doing ambulatory methodology, a persistent problem of context will occur. While laboratory responses have controlled stimuli, how ambulatory stimuli are described is much more complex. People's reactions and their environment will need thick description to help explain how responses occur. Otherwise, inferences will likely be incorrect and meaningful responses will be lost in the large variability of ambulatory settings.

I hope this chapter presents a strong case for advocating for mixed methods of qualitative and quantitative research. I see data science aiming to "replace" traditional ethnographic design, and I hope this plethora of examples make it clear that the data of skin conductance provides little information alone, in the same way multiple choice surveys or time in store measures hardly touch the surface of describing a customer experience.

I also hope readers appreciate the effort and expertise it takes to understand an emotional experience. A simpler design default is to project our own assumptions of emotions on the world: "I too would be stressed waiting in line for the TSA." But I think there is great potential in breaking down and dissecting an emotional experience to its root reactions. We can begin to discover how customers react in ways we never thought about. We can begin to watch people with a lens focused on their emotional state and not just their behavior. And as experience design becomes more and more popular, I see this as a growing necessity. After all, "If you can't measure you it, you can't manage it" – Peter Drucker.

Chapter 4: How Thick Psychophysiology Helped the LEGO Group bring the Building Experience to the tablet

4.1 CASE STUDY

My layover flight landed early – maybe I can make Thanksgiving dinner after all. But the woman behind the American Airlines counter ignores my panic as she drones on the phone. Another employee steps behind the kiosk and finds me an open seat. “Unfortunately, we charge $75 if you want to change your ticket to an earlier time. I’m sorry.”

Why would American Airlines charge me to move onto an earlier, open flight? I cross American Airlines off my short list of companies I love. Waiting an extra three hours in an airport is a benign punishment compared to other’s trials with the airline industry. Shannon’s luggage was placed too close to the engine and it incinerated. Even after multiple phone calls, the airline refused to reimburse her for their mistake. United Airlines lost Phoebe, a 10-year-old girl during a layover. When the worried at-home parents called United, employees refused to help, claiming it was not their responsibility. The article “*7 Stories that Prove the Airlines Hate You*” and the blog “Flights from Hell” affirms my experience was not unique.

Being disconnected from your customers is not terminal. Take for example the LEGO Group’s desire to bring the building experience to the tablet. The LEGO instruction team was designing a new app for Technic. Technic sets are some of the hardest models and can be a serious challenge for children with new pieces, gears, levers, and interacting pieces. I was asked to help the instruction team learn how Technic and instructions could integrate with tablets.

At the LEGO headquarters, I asked the instruction team why they needed an app. Technic was one of the most complex building challenges, and tablets could help lower the entry point, but no one knew exactly how. Nothing could be addressed, including the app, until we had a diagnosis of how Technic itself could improve.

With the room full of LEGO employees, I demo the MOXO sensors. Wendy wears the sensor as I start talking about the emotions of instructions. The sensors project lines that jump up and down as Wendy emotionally reacts to my talk. The MOXO sensors measure skin conductance, a marker of psychophysiological arousal. When Wendy becomes, frustrated, or excited, we'll see jumps in the signal. "Hopefully, we will not see any frustration during this talk," I joke.

Wendy stands up and introduces herself. Her skin conductance jumps: having an entire audience watch a participant is a great way to create stress. The children we measure will also have these jumps in skin conductance, which will help mark arousing moments. After each session, I then ask participants how they felt at these jumps, helping me characterize the emotions while playing with Technic. With the MOXO sensors and thoughtful observations, we can diagnose what is frustrating about Technic.

But the technology is not enough. We need to observe children's natural emotional reactions. Companies often conduct research at their headquarters. Imagine the excitement of a 7-year-old boy about to tell LEGO designers how they might make a better toy. That child's emotions are nothing like a child bored on a Saturday morning, deciding to try out Technic for the first time. So instead of holding the study at LEGO headquarters, we will visit children in their home and present the Technic set as a gift they can play if they like. I summarize this method as, emotion-centered prototyping – measuring and observing real experiences.

To meet our first participant, I drive with my co-worker Alyssa Reese an hour away to Valhallas in our car filled to the ceiling with Technic LEGO models. I manage to stall only four times on roundabouts as I simultaneously learn the Denmark driving culture and how to drive a stick shift. Mason, an "almost 7" year old boy, answers the apartment door. His room is layered in LEGO sets of spaceships, ninjas, and race cars. In the middle of the floor, LEGO Batman and Joker pose in a brawl. I ask Mason if he would be comfortable wearing the MOXO sensor. "Will it turn me into a robot?" "Only if you say the secret password," I play along.

Figure 4.1 Mason Wears the MOXO Sensors.

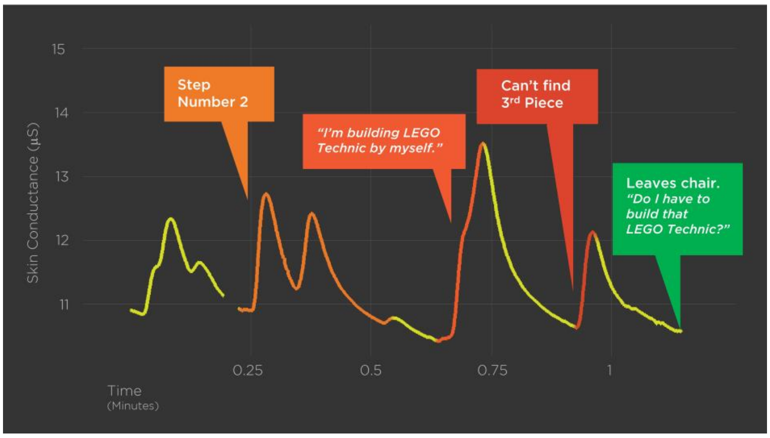


Figure 4.2 Mason is Stressed Building Technic.

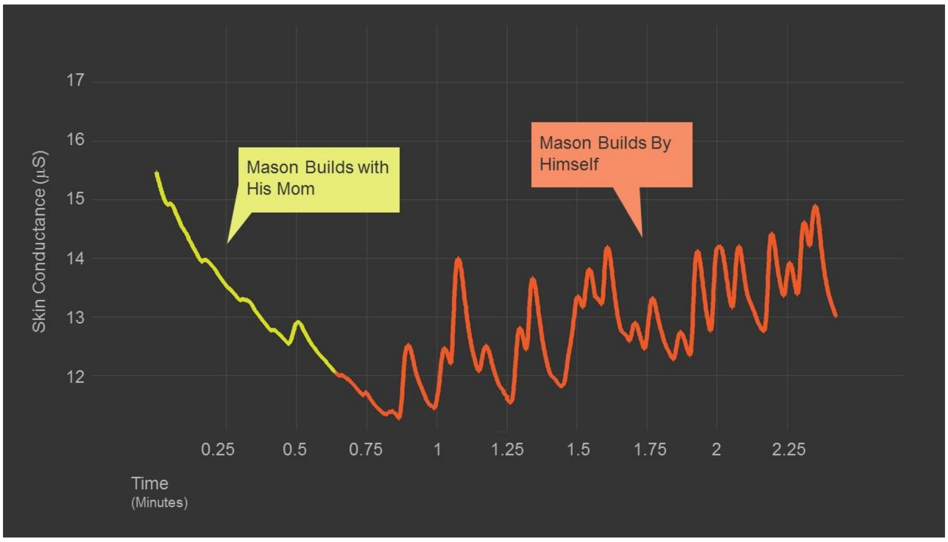


Figure 4.3 Mason Calms Down with his Mother.

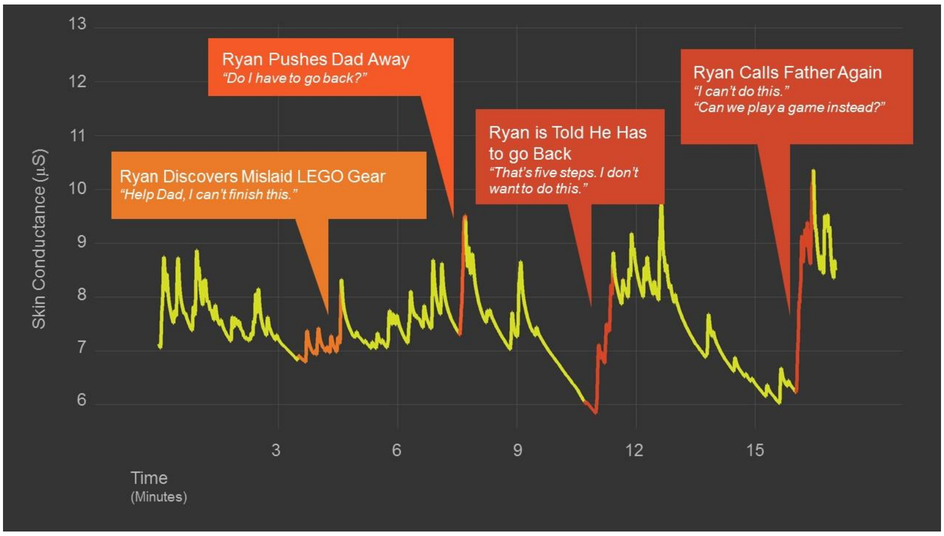


Figure 4.4 Ryan Stresses over a Misplaced Gear.

As Mason "waits for the sensors to calibrate," I offer him the Technic Back End Hoe. A movable shovel, wheels and hoe make this model more advanced than anything Mason has built. We leave the room and watch from the video camera. Mason picks up the first brick with eleven holes. He counts, "One, two, three, four, five, six, seven, eight, nine, ten, eleven." He then counts the holes in the instructions, "One, two, three, four, five, six, seven, eight, nine, ten, eleven." Step one complete. He turns the page, two more steps greet him. The MOXO sensors jumps. Making sense of so many new pieces is stressful and the new page is loaded with novel information.

"Mommy, I'm building LEGO on my own". A small bump from the MOXO sensor appears: Mason is a bit nervous. "I'm building LEGO Technic on my own." Mentioning Technic creates a much larger jump. He is now in a panic, he knows Technic is too advanced for him. Mason looks for the next three pieces, "One, Two. . . rup rup." He cannot find the third piece, the MOXO sensors respond again. Mason has found hundreds of LEGO pieces before. However, since Technic stresses Mason, even finding a third piece becomes a threatening task. A few seconds later, Mason walks to the room with his mom. "Um, Mommy, I have a question for you. Do I really have to build that LEGO Technic?" Mason quit Technic after three steps.

Later on, Mason's mother comes in and helps him build. The MOXO sensors show Mason calming down – his mother takes control and handles the difficult parts, helping Mason relax. I enter back into the room and watch Mason from my laptop. Hoping to impress us, Mason pushes his mother away. Mason's skin conductance starts increasing again. Mason pretends to drop a piece and crawls on all fours under the table, pops out at our end, and once more asks his mother if he has to build LEGO Technic. Building alone is stressful. I underline "parents are key" in my notes. This video will be the first time Marianne sees a child give up on building a LEGO model.

We later visit a 10-year-old, Technic expert, Ryan. His room hosts a Technic helicopter and car – models much more complicated than Mason's Batman set. Ryan shows us his iPad games like *Minecraft*. *Plants vs. Zombies* engrosses Ryan. I remind him of the free LEGO Technic model we have for him. Ryan reads the LEGO app sticker and replies, "I'll confer with the paper manual, but I really want to try out the app." Marianne and I sit in a neighboring room as Ryan's father sets down a plastic tray to hold the pieces. With the iPad on his side, Ryan starts off strong, building the first three steps with ease. However, on Step Four, he needs to insert a group of three pieces onto the larger model – a more advanced move. The orientation of the app instructions confuses him. He rotates the digital model back and forth a few times. Where exactly do the combo of pieces go? Ryan turns off the app and pulls out the instruction booklet, much simpler. Only four steps in and Ryan gives up on the app, the 3D rotating models are too complicated.

With paper instructions and forty minutes, Ryan's model nears completion: the body can move when he twists a knob. On step eighteen, he must place a brick over a gear, but the gear is attached to the wrong hole. Ryan calls to his father in Danish. The MOXO sensors jumps, asking for help is stressful. With some fiddling, the father removes the gear. Ryan now needs to put the gear in the right place so he calls the father back again. Marianne rapidly translates the conversation.

"Dad help again. I cannot finish. This cannot be finished."

"This is bad, I made a mistake."

"Dad, Dad, don't ruin it."

"You have to go back."

"Do I have to go back?"

The MOXO sensors responds again, and Ryan stresses over deconstructing his model. To cope, Ryan takes a hold of the model in an attempt to fix it. The father reaches for the model, but Ryan pulls it away.

“Did you skip something?”

“No Dad.”

“You have to go back.”

“That’s five steps; I don’t want to do that.”

Another jump occurs – going back is frustrating.

As Ryan attempts to fix the model, his father leaves him alone. In the moment of control, the MOXO sensors decrease.

After unsuccessfully placing the gear, Ryan calls back to his father one last time.

“Father, I can’t do this.”

“This is not fun.”

“I’m fed up.”

“Can we play a videogame together instead?”

The MOXO sensors show one final jump.

On a typical day, the LEGO building experience team works on problems about the number of pages, color format, and the ads in the back. Problems about introductions, reflections, and parents are new territory. Our work brought children’s frustration, anxiety, and excitement to the table. One employee told me how she could not sleep after the learning about the children’s struggles: so much needed to change.

Together we summarize the critical issues with the Technic experience. How children experience Technic for the first time is a vital step. How can we better transition from regular bricks to Technic bricks in a softer way? Parents need coaching as much as the child. What if we made a Technic kit specifically for pairs of parent and children? Ryan was stressed at having to go back in the instructions. Perhaps we could create ways for him to reflect on the model as he builds with animated gears? Once the LEGO Instruction team understood children's challenges, the purpose and vision of the app was a natural progression.

These questions expanded past the LEGO instruction team, and the research opened the gates for all the departments to help redefine what the building experience meant in the digital age. High level leadership and decision makers were invited in to discuss the future of LEGO on the tablet. Instructions that have been walled off and sacred for over twenty years were now open for innovation. I am excited to see how the LEGO group will reinvent the building experience, and what creative possibilities the tablet might unlock.

4.2 SCIENTIFIC ANALYSIS

4.2.1 Overview

Ten children's skin conductance was measured while they played with a difficult LEGO Technic set. In this section, we analyze two children's responses in detail. As parents differed how they interacted with their children, different skin conductance responses occurred. This preliminary study suggests that skin conductance responses could help identify more information about customers' psychophysiological arousal; however, researchers still will have challenges with specificity, and reliance on ethnographic and qualitative methods for interpretation are essential for inference.

4.2.2 Method:

Ten children, ages 6-12 were recruited for the study. A range of skill levels were included, in order to observe a variety of responses to the challenge of Technic LEGO. Observations were done in the children's house in places they would normally build. If more than one spot was suggested, we opted for a place with a table to keep physical movement more restrained. Children's skin conductance was measured from wireless Biopac sensors, referred to as MOXO sensors. Gelled electrodes were placed on the arches of both sides of children's feet. Preliminary tests on children's fingers showed a large number of artifacts due to their tiny fingers moving around. In order to create a real building experience, children were informed that the study was about sensors and not LEGO building. Preliminary tests showed children were reactive to the experimenter in the room – actively talking to the experimenter, showing their models, and mentioning their skill level. To help reduce the anxiety of having an observer, a webcam was setup with the observer sitting in the room next door[5].

After about one minute of the child sitting still, we told the children he looked bored, and we offered them LEGO Technic if they would like something to play with while we collected data on the sensors. We gently urged children to build the model for which we had tablet directions. Some children noticed the sticker that said tablet instructions, and for those children, we gave the tablet instructions right away. Otherwise, we offered the child a tablet 10-15 minutes after he had begun building, saying that a friend at the LEGO Group had new tablet instructions that could be used if wanted.

After the session, we asked children to tell us about the building experience, what they thought of Technic and the electronic instructions. For some children, if we had time, we had them review the video of them building and tell us what was emotional.

4.2.3 Results

We show the skin conductance of two children as case examples. Some of the 8 children showed similar responses and others different reactions, which is to be expected with children having different levels of expertise in Technic and parents using different teaching strategies. Further research would need to be done to understand how the responses shown would generalize to other children.

[5] Many children looked at the camera and had a response during this time. This suggests they still feel they are being observed. In future studies, it might be best to not label this as a LEGO study but as a sensor study, as children were still feeling the need to perform.

Participant 1, Mason

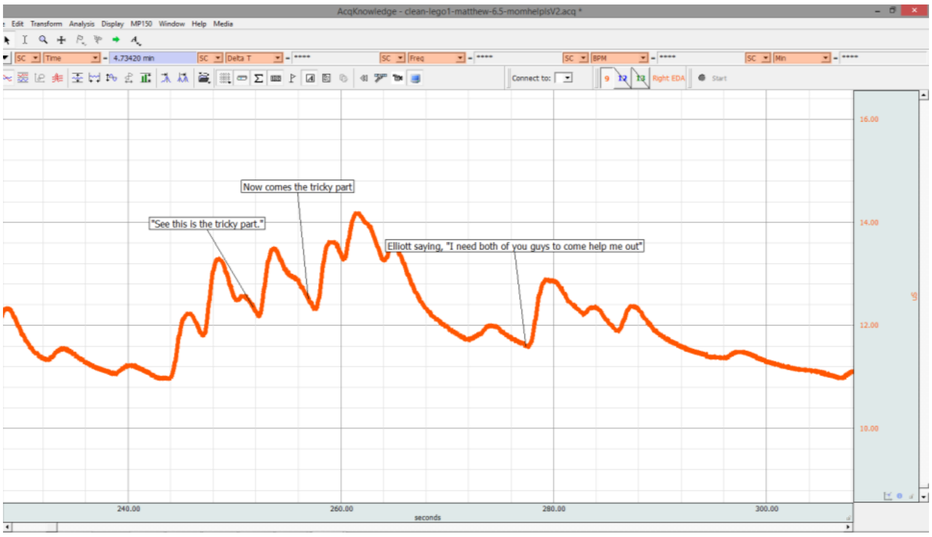


Figure 4.5 Mason's Skin conductance while researchers are in the room.

In the beginning, as we were setting up, Mason started talking to us. We can see two responses as he mentions that he is doing a tricky part. At this time he shifted his head in order to look at us from behind him. After this, we moved outside the house to make sure we were not distracting him.

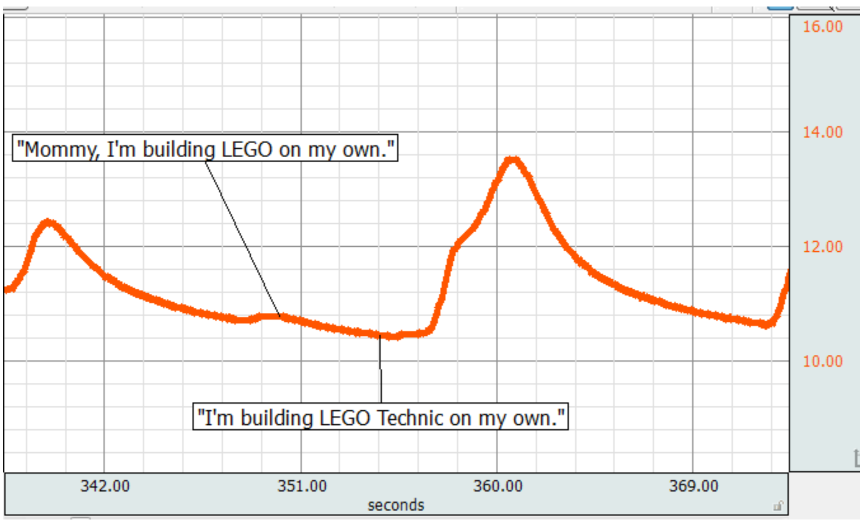


Figure 4.6 Mason's skin conductance talking about LEGO Technic.

As Mason begins building by himself, he calls out to his mother that he is building LEGO on his own. We see a small skin conductance response occur right before he talks to his mother, with an amplitude of 0.079 uS. This response is likely not due to his speaking, as a corresponding page turn occurred at this time. Nevertheless, Mason is talking to his mother with no large skin conductance response. Eight seconds later, Mason produces a much larger response saying the same phrase with the addition of the word Technic.

Mason talking in these two spots allows for a natural experiment. In both cases Mason is talking to his mother in the other room for about the same amount of time. The second response has a much higher skin conductance response. Since Mason's skin conductance only responded to the second time talking, the response was not likely due to the physical activity of talking itself. The second time Mason talks differs for two reasons. First, Mason mentions Technic. Additionally, this is the second time Mason mentions to his mother he is building on his own.

I conclude one of two options. First Mason may be pleading for his mother to help him, and the second response is larger because he is more worried. Alternatively, the concept of Technic affects him – either stresses him or excites him. Shortly after he quits building, I conclude that Mason is stressed by the concept of Technic. Interestingly, Mason appears to be smiling when talking about Technic. This smile may be an indicator that he was actually excited, but he might also be covering up his anxiety with a facade.

An interesting decision in interpretation is whether to measure the building LEGO technic response as one response with amplitude of 3.03 uS or two with amplitudes of 1.5 and 1.53 uS as the slope changes part way through. I have observed similar close responses many times, especially during stressful events. Visually, this appears as two responses, despite the slope never decreasing substantially (and in this case not at all). However, in interpretation, this halves the amplitude and might make the response comparable to a non-specific response elsewhere. Perhaps, a more appropriate way of classifying this second response is with response period. After the small peak turning a page and talking about building alone, another response does not occur for 8 seconds (7.5 responses/minute), with the next occurring when Mason starts talking about Technic. After Mason starts talking about LEGO Technic the second response occurs at 1.47 seconds (40.8 responses/minute).

A clear change occurs when Mason is interacting with his mother instead of by himself.

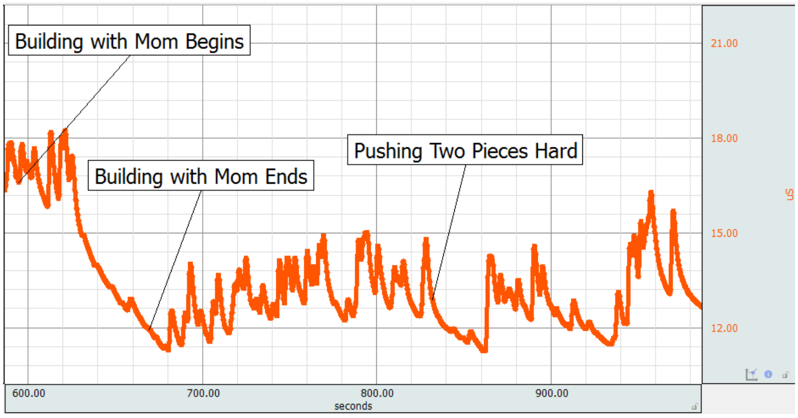


Figure 4.7 Mason building with his mother and alone.

As Mason and his mother begin finding pieces together, only one skin conductance response above 0.05 uS occurs, leading to skin conductance response frequency (SCR Frequency) of 1.4 responses a minute. An important note here is that skin conductance level gives a different story, with an elevated value 18.22 uS, despite the downward trend. This high value accompanies Mason walking over to his mother and asking for help. With fewer responses, we make the conclusion that while with his mother, Mathew's skin conductance response frequency is reduced. However, we can also conclude that Mason is also less likely to respond excitedly, and consequently may be bored. When we enter back into the room, Mason pushes his mother away and we see his skin conductance change dramatically. Working by himself, Mason's NS-SCR frequency changes to 8.61 responses per minute.

A challenge to interpreting this NS-SCR frequency is that Mason's actions also change when his mother leaves. He starts pushing the pieces together, which takes a lot of effort on his part. Likely some of these responses are due to physical movement, but are likely partially due to Mason's increased stress as well. An interesting comparison point is later on, when Mason pushes pieces together and does not show any skin conductance responses, suggesting that pushing pieces together may not always produce skin conductance responses. However, this may depend on the magnitude of force Mason uses as well.

Participant 2: Ryan

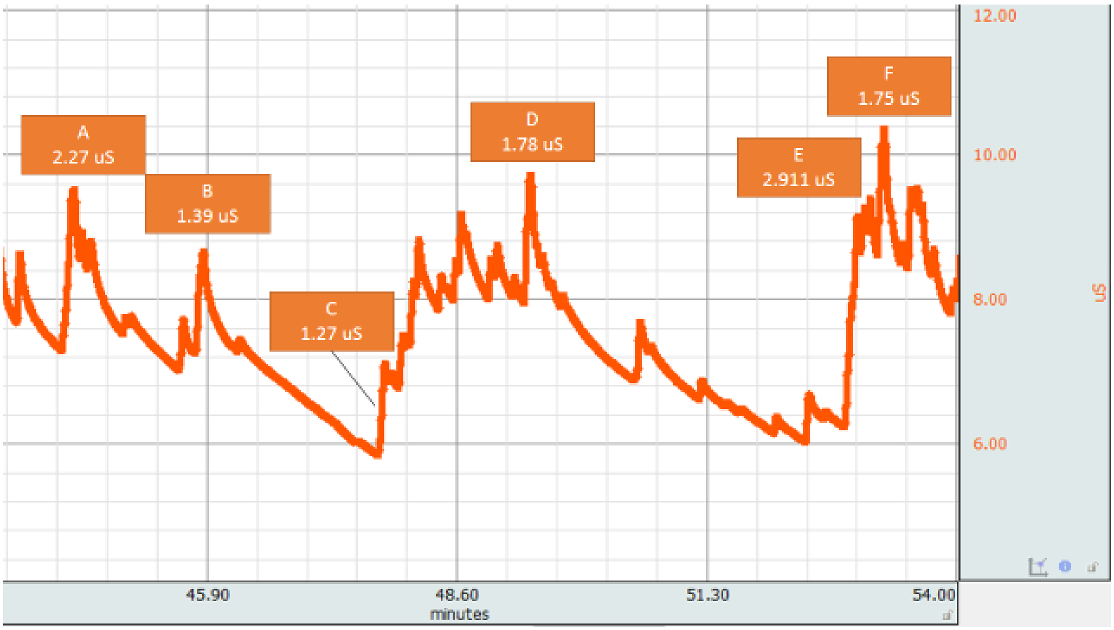


Figure 4.8 Largest Skin conductance Responses for Ryan during the end of the session.

Ryan's experience is complex as he reacted to 6 different stimuli across these 10 minutes. However, his response to these events were quite strong. These 10 minutes were selected because at the end Ryan began slapping his father and asking if he could play a video game again (an unwanted outcome of LEGO Technic). Following the suggested methods from Chapter 3, I first identified the responses with the top amplitudes, as labeled. Response B was removed from analysis as Ryan dropped a piece and bent over to pick the piece up. Response F was removed as Ryan was scooting his chair to protest building any further.

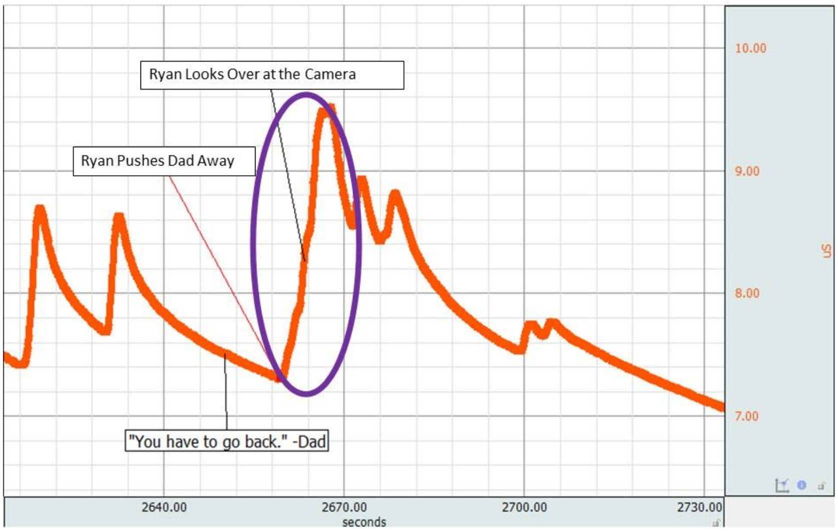


Figure 4.9 Response A in detail.

During Response A, Ryan is told he would have to go back several steps in the building instructions. He then gently touches his father, signaling him to move away and then looks at the camera. The response may be due to Ryan lifting his arm, but the large amplitude was likely due to more than just movement, as Ryan moves his arm throughout the 10 minutes. Since Ryan looks at the camera, I expect he is worried that we saw him push his father away or heard one of the researchers chuckle. While I expect this to be the main reason behind the response, Ryan may also be anxious about being told he has to go back, something that makes him respond highly in future sections as well.

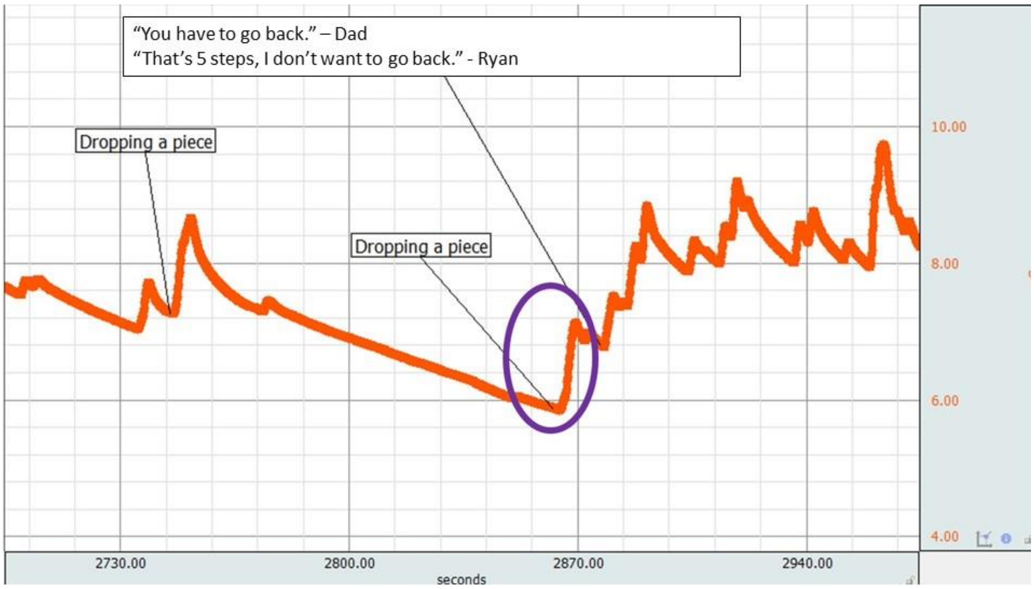


Figure 4.10 Response C in detail.

Response C has the lowest amplitude of the responses analyzed, but multiple responses occur closely after the initial response, suggesting Response C may be important. During Response C, Ryan drops a piece but does not lean back to pick it up; his father instead picks up the piece. Another response occurs after his father tells him he has to go back to an earlier step. The first response (Response C) I interpret as a reaction to the dropping of a piece. I interpret the following responses as a reaction to having to go back. This interpretation is further supported by the fact that when Ryan dropped a piece earlier, he did not show a multitude of responses after.

Figure 4.11 Response D in detail.

In response D, Ryan peers closely at the tablet instructions trying to understand what his next placement will be. During this time, Ryan leans forward to examine the tablet more closely. While leaning likely did have an effect on the response, we can see moments earlier when Ryan looked at the instructions without leaning, and showed a similar response.

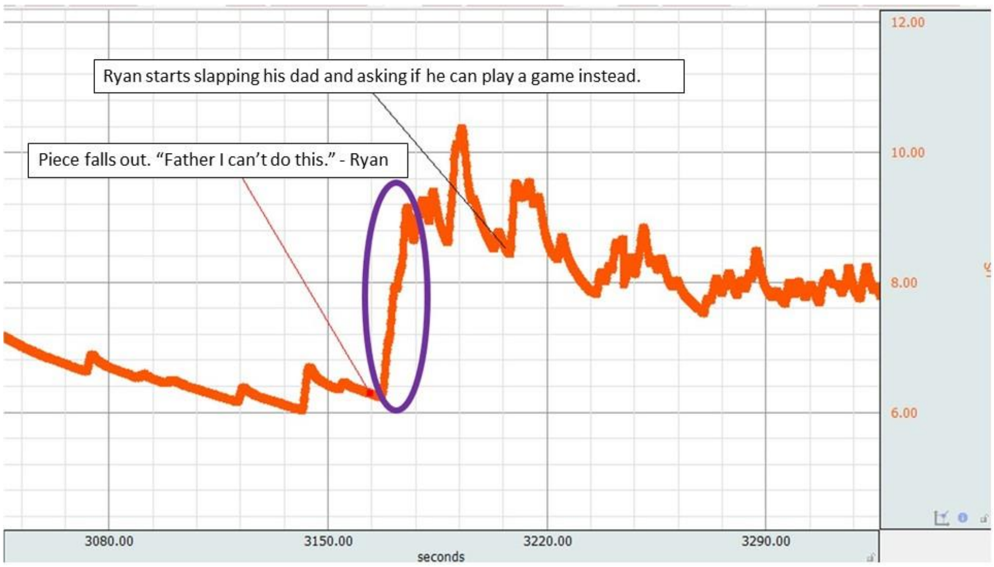


Figure 4.12 Response E in detail.

In this response, a piece falls out while Ryan is building, and Ryan tells his father he cannot do this, which in turn brings his father back to building with him. When his father starts building, Ryan starts slapping him on the back and asking if he can play again. I interpret this response as one of frustration and anger.

Conclusion:

This work is a preliminary attempt to show how skin conductance can help document the emotional experiences of children playing with LEGO bricks. Before we started this project, the LEGO instruction team did not view parents as a key predictor of children's emotional response to the challenge of Technic LEGO. However, our preliminary results suggested parents can have a large role in the emotional arousal of children. Both Mason and Ryan had large skin conductance responses calling out to their parents. And with parents there were moments when the children's SCR frequency decreased substantially.

However, the role of parents is not always equated to reducing stress. Mason's skin conductance also increased when his mother was doing all the work, and Mason appeared bored: he would tap pieces repeatedly on the table and play with one of the sensors on his hand (the data analyzed was from the foot). Ryan's skin conductance response increased substantially when his father told him he would have to go back several steps, almost as a sign of failure.

We discussed various other interpretations of these responses as well. By measuring skin conductance, we were able to see potential times when salient events were occurring. Future work should begin to look more closely at parent-children interactions. Perhaps a controlled study with specific instructions for parents could help us understand how parents can best keep children calm but engaged during building. Of course, we will still have the issue of children dropping pieces, giving varying pressure to the LEGO bricks, and being conscious of the researchers, but so is the nature of ambulatory research.

Chapter 5: How Thick Psychophysiology Helped the New World Symphony Engage New Concert Attendees

5.1 CASE STUDY

The people we serve experience the world differently than us. The CEO of Best Buy does not browse electronics like a grandmother buying her first computer. The engineers at Google use Google maps differently than a lost tourist, and conductors do not hear music the same way as their audiences. This divide in perception ruins experiences and chases away customers. But empathetic design research can bring these producers back to the customer. In this case, I explore how the classical symphony has shutout the majority of its potential customers and how it could respond differently to engage them.

> "I think I should have no other mortal wants, if I could always have plenty of music. It seems to infuse strength into my limbs and ideas into my brain. Life seems to go on without effort, when I am filled with music."
>
> – George Elliott, 1860

When I was a young pianist, the required classical symphony was torture. I would enter into the local church: white walls, spotless marble floors, and a stale air, frozen in time. An unknown pianist would sit down and play an unknown song by an equally unknown composer. It was a Concerto Number 6 or Symphony Number 7; it does not matter. At first, I was enchanted. The music was so complex and so different than my Red Hot Chili Peppers, but this engagement only lasted for a few seconds.

The notes quickly became the same –complex but nothing new, nothing to latch on to. The wooden seats kept my back unnaturally upright, squashing any hope of leaning back. I am sure this was intentional as I would have quickly fallen asleep otherwise. My mind would wander; I would look out the window at the leaves shaking in the wind – it was not much. I smuggled in a pen and doodle on my program. Twenty minutes later, the program would be filled with scratchings of aliens and astronauts, but the pianist was still only in the first movement. I would look at my watch, at least one more hour of this repetitive tune. The trees in the window had not moved. I would cross my legs and fold my arms and stare blankly towards the piano – it was not enough. On my side, my father had found a way to fall asleep, lucky him. I loved playing the piano, but sitting through a concert was torture.

Growing up, classical concerts did not "infuse my limbs with strength" or "make life effortless." I did not feel the humanity and emotional stirs that move the other attendees. Symphonies pushed me away from classical music – Beethoven and Bach were reserved for those over 65. I am not the only one who has been shut out; the number of people attending a classical concert has declined by 29% a year since 1982. Concert halls are closing down and shutting their doors, unable to find patrons and financial support. The National Endowment of the Arts estimates 234 million people in the United States never go to a classical concert. Article after article announce the end: "Is Classical Music Dying?" – New York Times, "Classic Collapse" – the American Spectator, and "Is Classical Music Broken?" – The Piano Society. The emotional pull of the classical concert has almost flickered out.

Looking for ways to reverse this trend, the New World Symphony asked me to help measure the emotional experience of a classical concert. At first I was hesitant. I was not one of the elite experts. I preferred Red Hot Chili Peppers concerts to *Hayden's Symphony No. 79*, but I was assured my outside view was ideal. After a few conversations, we narrowed our question: what could make classical concerts engaging again?

As a preliminary research study, I attended seven concerts. Some performances were strictly traditional while others attempted new solutions, including videos and emcees. From a list of ticket buyers, we selected 30 participants. A few participants were concert experts – attending the New World Symphony often. For others, this would be their first time at New World Symphony. We asked participants to join a long interview afterwards where they would re-describe their expectations and experience – what engaged them and what put them to sleep. And to understand their experience in a more precise, objective way, participants would wear the MOXO sensor.

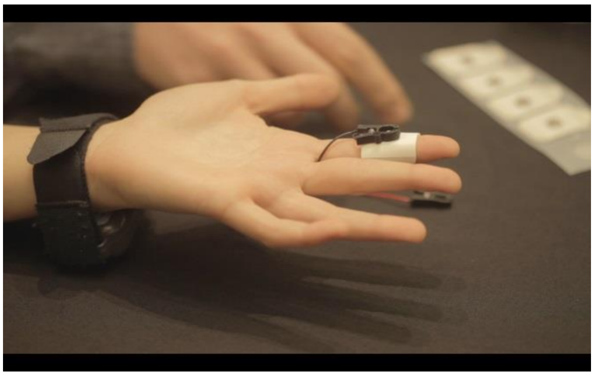


Figure 5.1 The MOXO sensor measures audience member's skin conductance. Skin conductance is a marker of the sympathetic nervous system, commonly known as the fight or flight response.

At MIT we have been working on ways to measure and understand people's emotional experience in the real world. The MOXO sensor measures skin conductance, micro-changes in an individual's sweat glands. When people become engaged or attentive, their skin conductance increases, and the sensor maps these changes. When large jumps occur, subjects are likely engaging with the music, being reenergized, and connecting with the world. The more jumps, the more reactive individuals are during the symphony. This sensor allows us to pinpoint times audiences react to the music. What in a symphony pulls at the heart strings?

Across the 30 individuals, skin conductance responses differed dramatically. Depending on an individual's exposure to classical music, we saw different responses. For participants who were not "your typical symphony goer" and had only been to one or two concerts, music often failed to grab them, especially during traditional concerts. Like me, attendees struggled to stay focused and wished the concert was shorter. "That last piece was quite long. Who sits there and composes this music for 40 minutes." "I get lost in the strings." Some of our participants even dozed off. "*The Firebird* is long, and it was the last part of the show, so it could be that I was tired and it lost me a couple of times." When asked to explain their experience, one attendee referred to baseball, an intimate pastime, "Someone who knows what is going on would like to see a 1-0 game, but someone who doesn't know baseball wants to see a slugfest." The MOXO sensor showed participants beginning energized but then drifting off, and eventually disengaging. For new attendees, *The Firebird* by Stravinsky was not a slugfest.

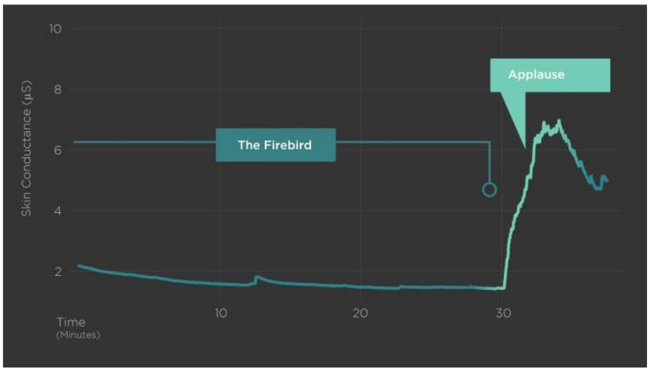


Figure 5.2 Skin Conductance Response of a Novice Listener.

For many new attendees to the symphony, classical music is difficult to connect with. This individual was unable to engage with *The Firebird* and correspondingly, the MOXO sensor showed few points of response (flat line). Afterwards the participant reported drifting off and wished the concert was shorter.

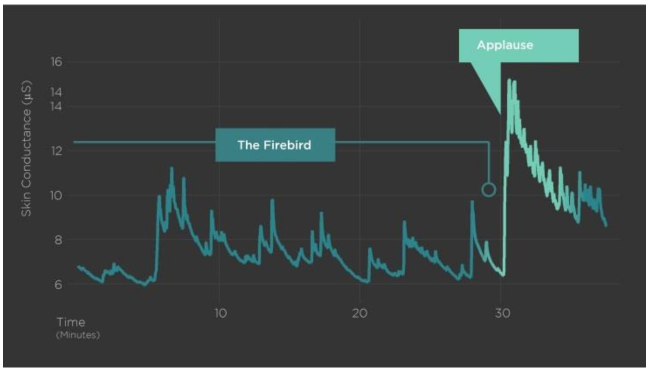


Figure 5.3 Skin Conductance Response of an Expert Listener.

For concert experts, the music itself is engaging, helping them stay alert and active. While listening to *The Firebird* this individual's skin jumped multiple times. Because he was able to react to the change in the music, he did not struggle to pay attention. During the interview he responded positively and looked forward to the next concert.

Other participants who were experts in classical music had a different experience. Michael, a staff member of the New World Symphony asked me after one concert, "Wasn't *The Firebird* amazing. I love that song every time I hear it, and tonight they nailed it." I chuckled – not knowing how to best explain my inability to connect to such complex music. I was hired to consult on classical music but unable to have that innate emotional experience the expert listeners were having.

Don was another one of those experts, listening to classical music every night. He told us how he fell into *The Firebird* and was swept up and down with the sweeping changes of volume. Looking at the MOXO sensors, we could see Don, like many other experts, react vividly to the crescendos, key changes, and solos, all grabbing his attention. Don connected with the music and that energized him.

We found a divide: experts connect to the music and the music gives them energy, which is not the case for novices. For that weekly concert-goer who can listen to every Baroque composer, symphony music is an emotional treat, something deep to react to. Nothing is broken or dying, nothing needs to be fixed. The experience just needs to be preserved. On the other hand, for the 234 million non-experts, classical music is not engaging. Going to a concert for the first time is like watching a movie in another language with no subtitles or watching a sports game without knowing the rules. Even worse, there is also no talking, no popcorn, and you cannot order a hot dog.

But symphonies are not doomed to be boring for novice listeners. If hot dogs and subtitles can fix entertainment, surely there are fixes to the symphony. We looked closer at the novice attendees' MOXO data and interviews and found three ways a symphony could better engage them.

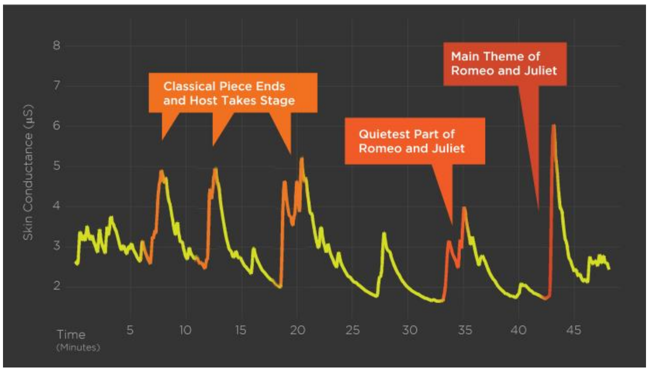


Figure 5.4 A Novice Listener Responds to "Symphony with a Splash."

While traditional concerts were hard for novices to react to, attendees were able to react to parts of Symphony with a Splash. This participant reacted whenever the emcee spoke or when the symphony performed *The Love Theme of Romeo and Juliette* accompanied by video.

5.1.1 Focus on the Familiar

Music we are already tied to and know will create emotional reactions. I opt for the John William's concerts when *Star Wars*, *E.T.*, and *Harry Potter* are performed – music with stories behind them. Every time *The Imperial March* plays, I tense up, imagining myself on the Death Star. *The Love Theme of Romeo and Juliette* brings back memories of commercials about falling in love. Participants praised *Hungarian Rhapsody No. 2* as it was the theme song of *Tom and Jerry*, a childhood pastime. Classical music would not be dying if Beethoven was a regular guest on Oprah.

In addition to playing film music, conductors can add stories to classics. Before Brahm's *Opus No. 6*, the emcee explained the history of the piece: Brahms was in love with Agathe but could not be with her, so instead sent her this romantic piece for her to play. Participants loved this piano solo; they could feel the emotion Brahms felt.

5.1.2 Transition Often

No matter the activity, with enough time, we become bored. Our need for novelty is why television shows stop after an hour and why there are three rings in a circus. Consistently, when a song ends and the emcee takes the stage, MOXO sensors show jumps–something new happened and the audience reacts. The ending of a song and transition to an emcee energizes us.

The emcee alone is not the reason for reengagement. In fact, when talks meander on, discussing a composer's political life or affair, audience members disengage, just like in the music. People react to the emcee as they are responding to a change in environment, from music to talk. The shorter the pieces are and the more clapping and speaking between, the more the audience reenergizes.

A conductor can create additional transitions in the music by highlighting key segments. For example, in the *Mother Goose* concert the audience was told a bassoon will begin to play, representing the beast in *Beauty and the Beast*. Participants then responded to this bassoon as it appeared in the music.

Figure 5.5 Audiences gather after the New World Symphony, sharing a drink and discussing the music. Moving around and socializing was a consistent high point in people's skin conductance.

5.1.3 Bring in Other Senses

During large rock concerts, a sizable investment is made in visuals. Having trouble focusing on the music? Watch the 40 foot movie. As the crowd shifts back and forth, constant sensations of movement and touch are generated. Mosh pits are incapable of being boring. As the concert plays, audiences can order beer and food as they please. The modern rock concert triggers more sensations than just sound and correspondingly has little trouble attracting new audiences.

The movie *Fantasia* has the same power to engage people through visuals. Not even in school yet, my brother and I would sit and listen to *Dance of the Hours* and other classical pieces for a full 90 minutes, multiple times. The difference between *Fantasia* and symphonies was a visual story that helped keep us focused and entertained. Hippos wearing pink tutus could be the key to making classical music engaging.

The Symphony with a Splash concerts had a visual supplement along with the music. When the Beast section begins in *Mother Goose*, a picture of the large projection of the boar in the princess' bed chambers fades in. When *Romeo and Juliette* was played, old black and white photos of the lovers hovered above the brass section. Even the slightest animation of leaves falling or subtle color changes provided a way for listeners to visually focus. "You are feeling the music and stuff but if you have something to see – that's really cool."

5.1.4 What will it Really Take?

With these early findings, reviving the symphony will be easy, right? Perhaps replace Beethoven with John Williams and add a giant movie—maybe one with lasers. But classical experts will be opposed, perhaps even irate, by these suggestions. Novices will insist on additional stimuli to connect to the music and experts will request the opposite, fewer distractions. The bells and whistles of movies, emcees, and modern culture can distract expert listeners from the core emotional pull of the music. While there are many, many more novices, the experts will likely determine the future of the symphony; they donate the money, attend the concerts, and dictate design. Unless symphonies focus on a larger, currently absent audience, I do not see concerts becoming engaging for the masses.

However, some venues have begun to reach out to those millions who could learn to love classical music. I look forward to this coming reformation– perhaps, one day I too will become a concert enthusiast. For those reformers, my one suggestion is to design from scratch. Do not tweak the symphony that is owned by the experts. Make something new, something dedicated to the novices. At the end of these concerts, do not ask your longest patron her opinion. Hunt down that new couple and hear their story. When they tell you it was too long, don't scoff; the concert was too long for them. Rather than blame the couples disinterest on their lack of classical upbringing, blame yourself as the designer. You and all your years of training and perfection will need to adapt in order to meet the audience where they are at—outside your doors, waiting to be infused with strength in their limbs and ideas in their brains. Waiting for life to go on without effort, waiting to love the symphony.

5.2 SCIENTIFIC ANALYSIS

5.2.1 Abstract

A total of 40 participant's skin conductance was measured during 1 of 9 varying classical concerts. After viewing the concert, subjects were interviewed about their experience and asked to describe their experience during moments of large skin conductance responses (over 0.05 uS in amplitude). Participants were divided into two groups, experts and novices, based on the number of concerts they attended. We compared an expert listener's response to a novice listener's. The expert had a large number of skin conductance responses during the performance, suggesting he physiologically responded to the concert. The novice had only one small response and did not physiologically respond to the classical piece. The novice verbally reported being bored while the expert reported enjoying the piece. In the

second analysis, we analyze the skin conductance response of a novice individual during a full classical concert. While she did not react during most of the music, we did see responses during transitions between songs and to *The Love theme of Romeo and Juliette*, a song with which she was familiar. These preliminary findings suggest that skin conductance response could be used as a way to detect when and how people react to music.

5.2.2 Method

Participants were recruited in pairs from people who were already signed up to attend the concert. Novices were selected from people who reported they had not attended a concert before. Participants arrived 30 minutes early and wore either a modified Q sensor or Biopac Wireless EDA Sensor with gelled electrodes on the non-dominant hand middle phalanx. After observing effects during clapping in the first session, participants were asked not to clap or talk to one another during intermissions as a means to reduce the number of non-emotional skin conductance responses.

Participants viewed either a traditional concert with an intermission or a “Symphony with a Splash” concert. The “Symphony with a Splash” concert was designed for novice listeners. These concerts were only an hour long, had Jamie Bernstein as an emcee, had several short music pieces, and included visual projections during the talk and during the performances.

After the concert participants were asked to describe how the concert went. Participant’s skin conductance responses were synced with video of the concert. Either skin conductance responses were visually identified or an algorithm identified moments of skin conductance responses. We replayed parts of the video where skin conductance responses occurred (starting 5 seconds earlier) and asked participants how they remembered feeling at those times.

5.2.3 Results

We show three example skin conductance responses as case examples. The first two subjects are the two subjects that we interviewed during a typical concert. We measured other people’s responses, but I learned afterwards from the interviews that expert and novice concert attendees were incorrectly assigned. Many of the labeled novices were in fact experts who had not attended the New World Symphony before, but had attended multiple concerts elsewhere. Because only two regular concerts were measured, and the definition of expert shifted with preliminary interviews, a follow up study should take place where 8 novices and

8 experts listen to the same concert simultaneously in order to understand how differentiable expert and novice responses are in general. The third subject was selected as she was a novice and heard a song with which she was familiar during the concert, demonstrating a key finding.

Subject 1

Subject 1 is an expert classical listener, who listens to classical music every day and reported enjoying going to the concert. He attended a traditional concert. The song after the intermission was *The Firebird* and was conducted by Esa-Pekka Salonen a world famous composer and conductor from England. The participant reported enjoying the concert and was classified as an expert. Below is his skin conductance response for the last 30 minutes of *The Firebird*.

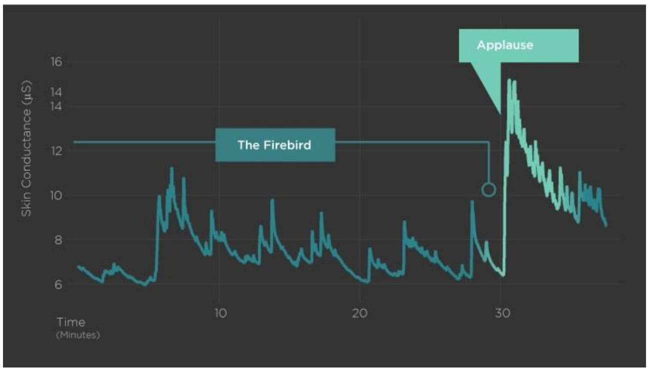


Figure 5.6 Subject 1's Skin Conductance Response to the last 30 minutes of *The Firebird.*

His skin conductance response frequency (SCR Frequency) was 3.76 responses per minute. (Responses include all responses over 0.05 uS.) In order to show these responses were likely from the environment and not just a trait characteristic, we can show times of non-reaction too. Here is the response of that same individual for the entire song.

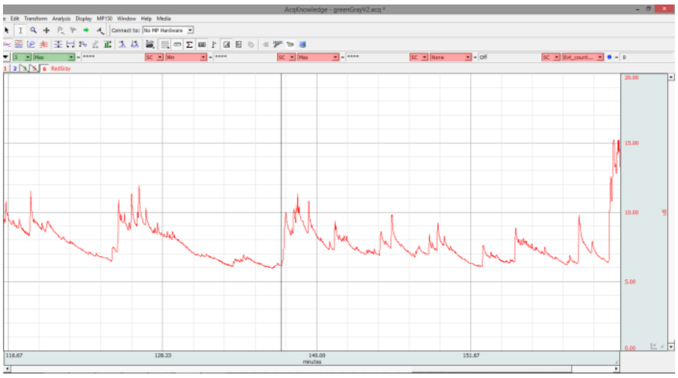


Figure 5.7 Subject 1's Skin conductance shows less skin conductance responses while listening to an early part of the concert.

Notice how earlier there were fewer skin conductance responses, suggesting this subject's skin conductance response frequency is not always high.

Subject 2

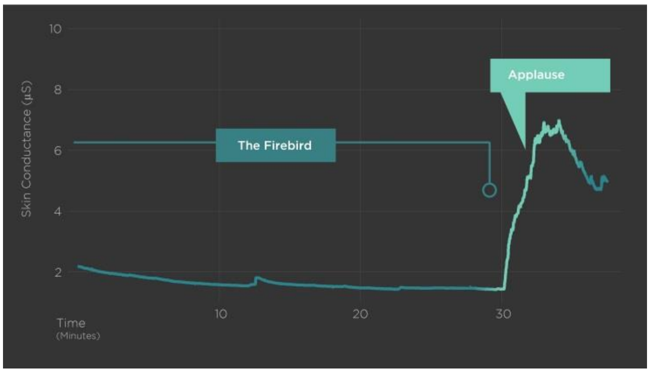


Figure 5.8 Subject 2's skin conductance response during the last 30 minutes of *The Firebird.*

Subject 2, a novice listener shows only one small response at 12 minutes, with a SCR frequency of 0.038 responses per minute. In order to show that this individual is capable of having skin conductance responses, one could look at his skin conductance at the beginning of the performance when a short introduction video was played. At this time, the individual's skin conductance response was 2.09 responses per minute.

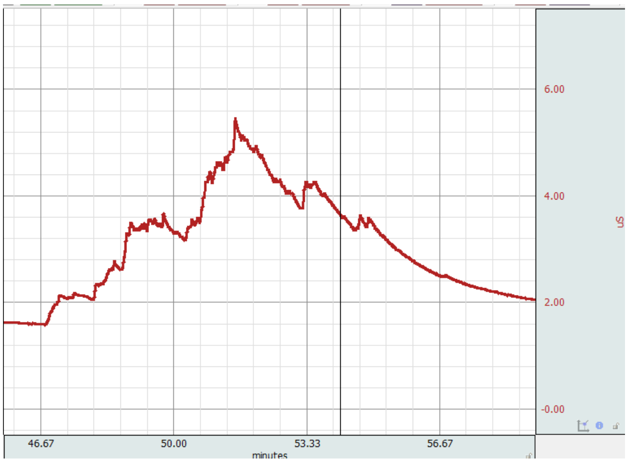


Figure 5.9 A novice listener shows skin conductance responses at the beginning of the video.

Ideally, we would measure more novice participant's reactions to the end of The Firebird to determine if this non-response was truly a non-response or if the individual transitioned from being labile to stabile. Most novice participants showed non-responsiveness to long periods of music, so I do not believe this is a case of the participant being stabile, but analyzing multiple subjects' responses simultaneously would help substantiate that belief.

Below is an estimate of five people's skin conductance response frequency listening to the same concert from one of two nights, during non-clapping sections. Researchers could collect more data during similar concerts to see if SCR frequency statistically differs between novices and experts listening to traditional concerts.

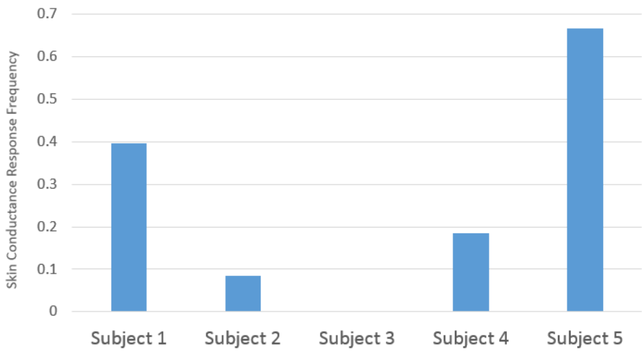


Figure 5.10 Skin conductance Response Frequency of 5 participants listening to the same concert.

Subject 3

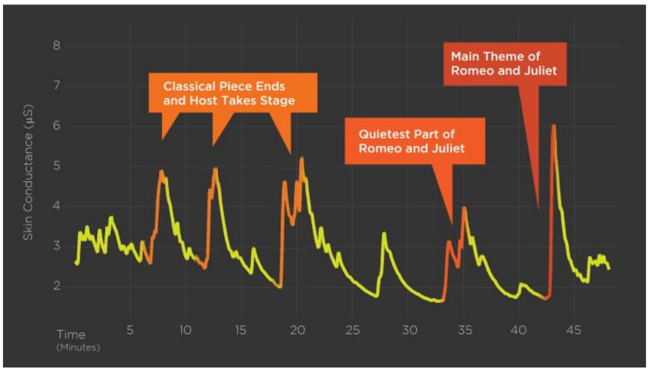


Figure 5.11 Subject 3 is a novice who reacts to various parts of the concert

Subject 3's responses illustrate what novices might react to. The participant had never been to a classical concert but was excited about the idea of Romeo and Juliette as a theme. She reported enjoying the music and said she planned on buying a ticket to a future show. We see large skin conductance responses during transitions when the music ends, the participant claps, and an emcee takes the stage to speak with a visualization in the background. These large responses were reliably found across participants. When participants were instructed not to clap or talk to their partner at these moments, we still saw large jumps, suggesting song transitions increase physiological arousal. We can also see that the emcee's talk was not the source of arousal as the participant's skin conductance decreases afterwards.

While the participant did not typically react to the music, two large responses occurred during *Romeo and Juliette*. One was during the quietest part and the other was the main theme. All four participants that night had a response to that song; see (Hedman, 2013) for a detailed analysis. These two sections were reported to be emotionally moving and familiar. The song is a popular song often heard in commercials and movies and participant 3 reported being familiar with the song as well.

5.2.4 Discussion:

While this study was primarily conducted as an exploration of the usefulness of skin conductance sensors, we provide two questions that could be explored more.

Question 1: Is Low Skin Conductance a bad state?

Expert participants reported wishing there were fewer distractions so they could "just listen to the music." Multiple experts reported the "Symphony with the Splash" songs were too short, ending when they were just getting into the music, suggesting a low arousal state may be needed to connect fully with the music. Conversely, novice listeners reported liking the emcee and visuals as it “engaged” them. Therefore, it appears low arousal has its benefits and consequences. From a benefit side, no distractions means there is nothing else to react too, and one can listen to the music peacefully. However, as arousal decreases, it can become harder to concentrate and focus (Thayer, 1989). Perhaps the visual aids give novices something to focus on more easily. These findings suggest that novices in a low arousal state need something easy to focus on or they get distracted and have to struggle to engage and that is the feeling of boredom.

Question 2: What are some ways to increase arousal?

Transitions appeared as a strong way to increase arousal, with participant’s largest skin conductance responses occurring during the ending of songs. Many interviews turned into us just showing one ending after another. Because of this finding, we recommended the New World Symphony keep songs short in “Symphony with a Splash” to allow novices a break from that low arousal state. An interesting question is what parts of the transitions are creating that jump in arousal? If the symphony went from one movement to another would this still occur? Or is it the clapping sound? This becomes an interesting question, as clapping has been disallowed between movements in recent years and has been debated about whether it should come back. Alternatively, is the fact that an emcee is coming on stage a key way of creating these responses?

Familiarity is a much less tested one, as *Romeo and Juliette* was the only song novices reported knowing. However, previous music literature points towards the importance of familiarity (North & Hargreaves, 1997). Future work should test this hypothesis at a John Williams concert or other places where familiar movie works are played. This theory is supported by research which suggests skin conductance responses increase with familiar stimuli (Critchley, 2002a; Morris et al., 2008).

5.2.5 Conclusion

Measuring skin conductance helped designers understand the classical concert in ways traditional interviews could not. No participant we interviewed reported “liking transitions” but this response was a consistent moment of high arousal. Similarly, asking people whether they “reacted” to the music is difficult – how does one gauge and remember their reactivity? Our initial data showed strong differences between some expert and novice listeners. Future work should look if these differences are a generalizable trend and also explore other ways people engage with passive media.

Chapter 6: Psychophysiology as an Effective Boundary Object

6.1 BACKGROUND

Together, with the LEGO instruction team, we measured the skin conductance of 9 children as they played with a challenging LEGO Technic set, as a form of empathetic design (Leonard & Rayport, 1997). One key insight was the importance of parents being present during the challenging portions of building: a parent could help reduce the stress of building a LEGO model if invited in (see Chapter 4). From a design research perspective, this insight was ideal. The LEGO instructions team had a whole new, unstudied space to explore: including parents in play. By better including parents into play could help increase the sales from selective mothers, help children transition to more difficult sets, and increase the play value of the sets.

Learning about the importance of parents was the easier challenge. The harder challenge was transferring this knowledge to the LEGO group and motivating teams to focus more on the parent experience. We sent an initial survey, asking LEGO employees how they felt about an MIT student researching ways to build “Mom-Friendly LEGO”. Multiple employees replied back that LEGO bricks were already “Mom-friendly” and research should work on being more “kid-friendly”. Different departments across the organization asked why they were not included in the initial framing of the project. Stakeholders pointed out that the case child was 6 years old, while the Technic set he was using was prescribed for children 7 and older. Reacting to these disgruntled comments, we changed the ending of the survey to re-emphasize that these initiatives were hypothetical and not being executed. These comments were not surprising, spreading knowledge to different groups in an organization is difficult. However, sharing knowledge across domains is also a source of competitive advantage; innovation is mostly likely to occur at these boundaries between disciplines (Leonard, 1995).

Dubberly suggests that transferring design knowledge could be worth billions of dollars (2012). He gives the example of Samsung’s chairman, Kun-hee Lee, who understood the value of design and put substantial resources into creating a top-tier design company (Freeze & Chung, 2008). And yet, despite the considerable investment, Samsung has failed to achieve the same design level as Apple. At the same time Sony and IBM, both previous design leaders have fallen into the background of the design world. While a great number of sources discuss design research methodology, much less research is available on communicating design research and creating design understanding within a company. No matter how great a design idea is, if it is not embraced within an organization, it will go to waste. Building momentum and support for human-centered ideas is core to design effectiveness, and may be the most difficult part.

A key way of transferring knowledge to different groups is with boundary objects. Boundary objects are artifacts or drawings that help multiple groups with different perspectives communicate ideas to one another. Star & Greismer (1989) describe boundary objects as "an analytic concept of those scientific objects which both inhabit several intersecting social worlds (…) and satisfy the informational requirements of each of them." The researchers give an example of how the classification of species at a museum is a boundary object between a professor at a museum and collectors. Knowledge at boundaries can be framed by three relational properties: difference, dependence, and novelty (Carlile & Rebentisch, 2003). Differences are defined as the degree to which two groups differ, typically in the degree of knowledge or specialties. Dependence arises when groups must rely on one another to complete a task. Novelty occurs when the task has new conditions or requirements, which is the most challenging property. Actors are susceptible to believing novel knowledge is already known (Martins & Kambil, 1999) or irrelevant (Arrow, 1994). Referred to as "the curse of knowledge," employees can struggle to abandon previous held beliefs and adopting new information that contradicts those current beliefs (Camerer et al., 1989). Bechky (2003) finds that knowledge-sharing difficulties are rooted in differences in language, locus of practice, and conceptualization of the product being designed. Carlile (2002) builds off of boundary objects and defines effective, pragmatic boundary objects as objects that facilitate a process for individuals to jointly transform their knowledge. These effective boundaries are essential for design research to be communicated.

Carlile gives the example of Mick working on a recovery valve within a 300-person business. As the sole manufacturing engineer at the early stage of a project, he noticed the complex valve needed to be divided into four subassemblies. He presented these findings on an out-of-date assembly drawing. When Mick went to the review meetings, his suggestions seemed to only anger the sales representatives and design engineers. This assembly drawing did not reflect the design team's own concerns – the design of the current valve.

With 8 weeks left, Mick once again presented the need for subassemblies to the workgroup, but this time with an updated assembly drawing, with correct tolerances, specs, and locations. While the argument was the same, the design engineers reacted differently to this proposal as Mick could reference parts that were "at stake" for the design engineers; they could see how Mick's changes would affect their areas of concern in the updated model. Mick gave the same message but communicated with a different object. Carlile suggests that some boundary objects are more effective at transformation knowledge than others and are context specific: CAD drawings do not work in all situations.

How does a researcher best share his or her tacit knowledge of the emotional experience of children playing with LEGO bricks? In the LEGO study, we hoped to share knowledge about the emotional experience of children. While CAD drawings and maps are good candidates for boundary objects, affective knowledge cannot be as easily and successfully placed in a diagram or object. Design insights are often subjective, customer experiences (Berger, 2014). Perhaps, this is why designers often use stories as a means of transferring insights. However, these qualitative stories are not always received well, especially with communities that mainly use quantitative measures (Diefenbach, 2009; Light, 1993). In response, some design firms have begun looking at including more quantitative data as a way of making findings more accessible (e.g. Seemann, 2012).

While the majority of pragmatic boundary objects studied have focused on maps and diagrams, I propose a new type of boundary object: psychophysiology. Psychophysiology can help communicate a person's emotional experience in a scientific, numerical way. In other fields outside of organizational management, psychophysiology has been shown to be a persuasive way to communicate psychological effects. A variety of studies have shown participants a "bogus pipeline" of fake electromyography or skin conductance (Jones & Sigall, 1971). Typically, in these studies, a psychophysiological device is "validated" in front of the user with researchers altering the feedback of the device behind the scenes so the participant believes the device can accurately measure their affect. While participants wear the devices, they are asked a series of personal questions and their answers are compared to controls' answers. Males who wore an EMG device self-reported African Americans as being less honest than controls wearing no device (Sigall & Page, 1971). Participants who were given false feedback of their heart rate reported faces as being more emotional (Gray et al., 2007). Heart rate false feedback has also led participants to report feeling more anxious and underperforming during a social task (Wild et al., 2008). False feedback about elevated heart rate has also led people to be more likely affected by a fearful communication (Harris & Jellison, 1971). Telling an individual false information about his or her psychophysiology can alter their own self-reported emotional state.

Self-reported affect is also altered when participants are shown psychophysiological data of others. Readers are more likely to agree with scientific articles if the articles have a brain image with them (McCabe & Castel, 2008). This agreement occurred with articles with faulty logic and sound logic. However, this increased agreement only occurred with brain images; graphs did not produce the same effect. Weisberg and colleagues (2008) found higher confidence with brain images did not occur for expert neuroscientists who were able to detect the logical flaws in the scientific articles presented. Both studies point to brain images as effective boundary objects for scientists to communicate to lay audiences. This increase in confidence of inaccurate but "scientific" findings has led to polygraph evidence being ruled inadmissible, as it will bias a juror's fact-finding role (McCabe et al., 2011). Beck (2010) reasons that an attraction to brain images reduces the complex process to overly simplified statements of "X causes Y". There may also be an appeal to biological responses as an untrained reader may take the data as innate, unchangeable traits about humans.

Empathy is an additional factor that can impact client's understanding and support for an initiative: do employees empathically respond to the stress or frustration of presented customers. Affective empathy can help increase a person's helping behavior (Batson et al., 1987; Eisenberg & Miller, 1987; Krebs, 1975; Toi & Batson, 1982). In addition to affective empathy, "responding with the same emotion to another person's emotion"; individuals can also have cognitive empathy, "intellectually taking the role or perspective of another person" (Gladstein, 1983). Cognitive empathy can alter how people attribute behavioral responses (Gould & Sigall, 1977; Regan & Totten, 1975). Neuroscientists have found a double dissociation between the cognitive and emotional empathetic responses (Shamay-Tsoory et al., 2009). However, multiple studies suggest that the cognitive and affective elements of empathy are inseparable (Feshbach, 1975; R. L. Katz, 1963; Schafer, 1959; Strayer, 1987). This study measures affective empathy as an additional factor that could influence employee's support for an initiative (Batson et al., 1987).

6.2 QUESTIONS

1. By presenting overlaid psychophysiological data of a customer, are LEGO employees more likely to increase the priority of an initiative compared to a traditional design story or a video without skin conductance?
2. Are employees more likely to increase the priority of a project if given pre-defined physiological data or open-ended, interpretive data?
3. How do displayed skin conductance, reported empathy, and increasing priority interact?

6.3 METHOD

143 employees from the LEGO group (72 females) completed an 8-minute, online questionnaire sent to their work email. Participants were randomly assigned to one of four experimental conditions: 26 in Story (S), 39 in Video condition (V), 49 in Defined Skin Conductance Responses (DS), and 37 in Open Skin Conductance Responses (OS). The four conditions differ on how participants were told about current research on the initiative. The S condition told a short story of when Matthew's mother helped him remain calm accompanied by a picture of Matthew. The V condition showed a video of Matthew struggling by himself and then his mother helping him calm down. The DS and OS showed the same video with skin conductance overlaid. In the DS condition, skin conductance was described as a marker of stress, and in the OS condition skin conductance was described as increasing for a variety of reasons like breathing, moving, and becoming stressed.

6.3.1 Introduction

Participants were emailed with the subject line of "LEGO directs the next phase of MIT student's research. In five minutes, help us decide how." After a brief explanation of the survey, participants read a description of the MOXO sensor, which measures skin conductance.

6.3.2 Measuring Importance

Participants were given a description about "Creating a Mom-Friendly LEGO Building Experience" and then asked "How high of a priority should 'Creating a Mom-Friendly LEGO Building Experience' be for the LEGO Group?" on a 5 point scale (1=Lowest Priority, 5=

Highest Priority). 16 of the 143 participants were removed from analysis, as they listed Mom-Friendly Building as the highest priority and could not increase that priority.

6.3.3 Describing the Science

Participants were then randomly assigned in to one of the four conditions. In the S, V, and DS conditions, participants were told “When an orange spike occurs, the child is more likely to be stressed.” Skin conductance responses were highlighted in orange. In the OS condition, participants were told “When a child becomes stressed, excited, thinks hard, breathes hard, or physically moves, the MOXO data will spike up (orange lines). You will need to look at contextual clues like what the child says and how he behaves to determine why spikes might have occurred.”

6.3.4 Giving a Case Example

As an argument for creating a Mom-Friendly LEGO building experience, participants were given the story of Matthew in four different ways. The S condition said researchers interviewed families, followed by a short story of Matthew with a picture of Matthew holding a LEGO box. The V condition said researchers videotaped 10 children, and a 45 second video of Matthew was shown. The video is described to participants as “Matthew builds his fist Technic alone, and the instructions become too hard. Soon after the video clip, Matthew stands up and asks his mother if he can quit playing. When his mother helps him, he stays calm.”

In the DS and OS condition participants are told that 10 children wear the MOXO sensors with the same description of the video as the V condition. The video is the same as the V condition with the exception that Skin Conductance is overlaid onto the video. Afterwards, all participants were asked to rate their priority of the initiative again.

Videos can be viewed here:

http://youtu.be/QOcUMcXd6F8

http://youtu.be/tYUKpXg-C8g

6.3.5 Change in Priority

After viewing the case example, participants were then asked again, “Now, how high of a priority should "Creating a Mom-Friendly LEGO Building Experience" be for the LEGO

Group?" on a 5 point scale (1=Lowest Priority, 5= Highest Priority). Participant's new rating was subtracted from their original rating to calculate their change in priority. If change in priority was positive, they were labeled as Grew Priority (GP).

6.3.6 Empathic measurement (Batson et al., 1987)

Participants described the empathic reaction to the story by rating a series of 6 adjectives (compassionate, tender, softhearted, sympathetic, moved, and warm) on a 7-point scale ranging from 1 (not at all) to (extremely). An additional 5 adjectives were placed in between the empathic adjectives. By averaging the values of the empathic adjectives, we created an empathy score. Following Batson and colleague's example, participants with the survey's median score below 4.0 were labeled as Low Empathy (LE) and those with scores 4.0 or higher were labeled as High Empathy (HE).

6.4 ANALYSIS

Graphed below is participant's change of priority for the Mom-Friendly Building initiative. This data was derived from subtracting the participant's initial reported priority for Mom-Friendly LEGO from his or her reported priority after viewing the case study. Participants labeled as GP had a score of 1 or 2 (a positive change).

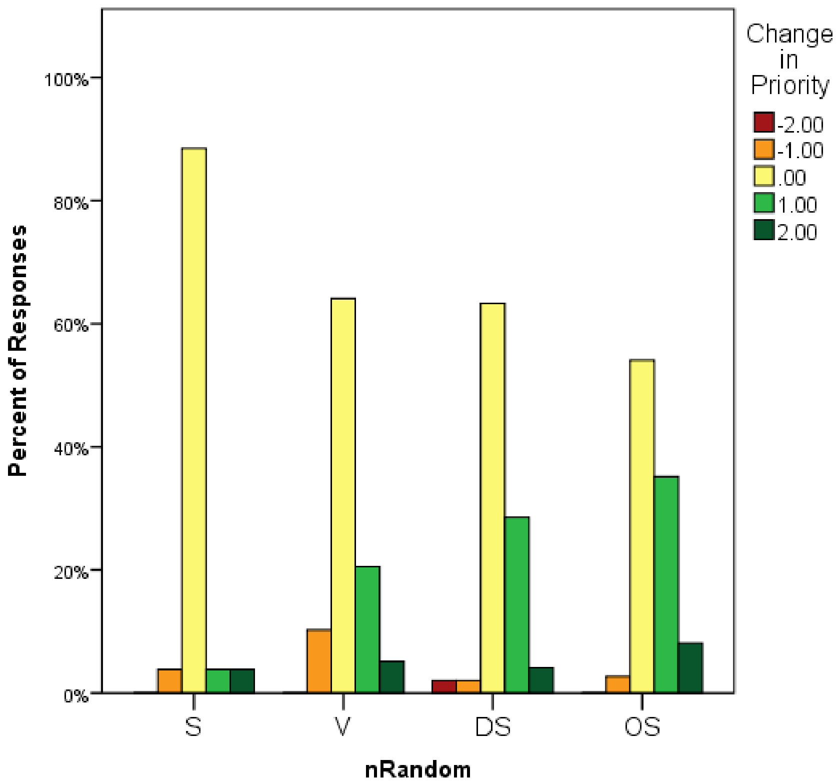


Figure 6.1 Distribution of Employees Change of Priority for the Mom-Friendly Building initiative (Positive Means an Increase in Support).

An analysis of variance with Condition as the independent factor and GP as the dependent factor showed a main effect of Condition ($F(1,3)=3.45, p=0.019$) indicating that conditions had an effect on whether participants increased priority for the initiative. A LOGIT model was applied to the data. The additional factors of sex, having boys, being a mother of boys, and initial priority rating were also tested (see appendix) but had no statistically significant interaction, so were removed from analysis.

Table 6.1 LOGIT Model of Condition's effect on GP with S as the reference condition.

	N	B	e^B	S.E. of e^B	Sig.
STORY (S)	23				
VIDEO (V)	35	1.435	4.2	2.2	0.084
DEFINED SKIN CONDUCTANCE (DS)	43	1.828	6.2	2.2	0.023
OPEN SKIN CONDUCTANCE (OS)	34	2.234	9.3	2.3	0.006

Both the OS and the DS condition were significantly different from the S condition ($p<0.01$ and $P<0.05$). OS and DS did not significantly differ, ($p=0.542$). There was a marginally significant difference between OS and V ($p=0.055$), suggesting that the OS condition's increase in % GP was due to more than the effect of a video. As an alternative method of analysis, a Bonfferoni test was done between conditions, with the only statistically significant difference being between OS and S with $p=0.014$.

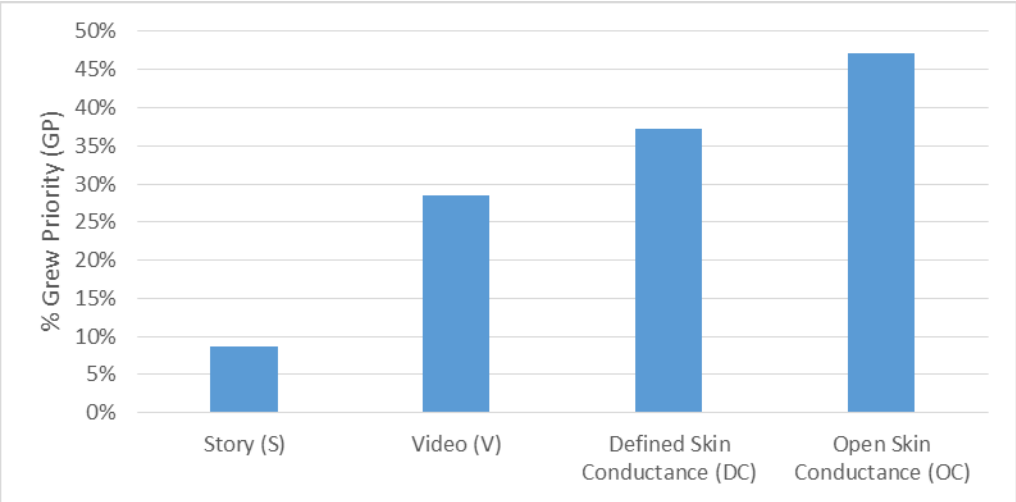


Figure 6.2 Participants were more likely to grow priority of an initiative after watching a video of skin conductance compared to reading the same story.

An analysis of variances with Condition as the independent factor and Empathy Level as the dependent variable did not show an effect ($F(3,109)=1.684$, $p=0.184$). However, empathy scores and GP showed a significant correlation ($r=0.338$, $p<0.001$).

An analysis of variance with Condition and Empathy Level as independent factors and GP as the dependent variable showed that both Condition ($F(3,119)= 4.174$, $P<0.01$) and Empathy score ($F(1,119)= 5.756$, $p<0.05$) had a main effect, with the interaction between the Condition and Empathy Level having an almost significant effect ($F(3,119) =2.2541$, $p=0.087$). Because 0 of the 13 participants in S (LE) respondents grew their priority (GP), we were unable to do a LOGIT analysis for LE compared to story. Within LE, the other 3 conditions did not differ between each other. A Bonferroni test between LE conditions showed no statistical significance. Below is the LOGIT analysis for the HE group.

Table 6.2 LOGIT Model of Condition's effect on GP for HE participants with S as the reference condition.

	N	B	e^B	S.E. of e^B	Sig.
STORY (S) (HE)	7				
VIDEO (V) (HE)	19	3.50	4.2	3.32	0.297
DEFINED SKIN CONDUCTANCE (DS) (HE)	21	2.92	6.2	3.30	0.370
OPEN SKIN CONDUCTANCE (OS) (HE)	12	15.00	9.3	3.60	0.035

OS (HE) statistically differed from S (HE) and DS (HE) ($p<0.05$) and OS (HE) approached significance with V (HE) ($p=0.079$). As an alternative analysis, a Bonferroni test showed a statistical difference between OS (HE) and S (HE) ($p=0.034$). The percentages of GP are graphed below for both empathy groups:

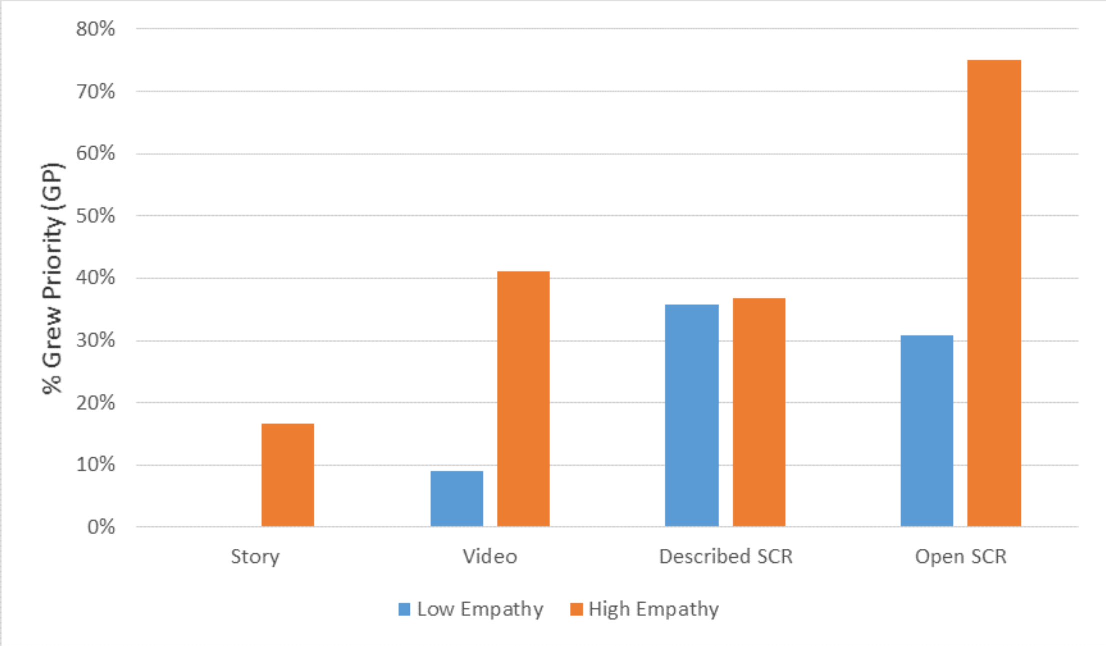


Figure 6.3 Participants likelihood to grow priority, split by self-reported empathy.

6.5 DISCUSSION

In this experiment we found that videos with skin conductance increased the percentage of people who increase the priority of an initiative compared to a traditional story. Even though the same information was presented in all four categories, psychophysiology and video were more effective at increasing priorities of an initiative. Participants with high empathy scores in the openly interpreted skin conductance (OS) condition were even more likely grow the initiatives priority (GP), whereas participants who were told skin conductance only equates to stress (DS) were not more likely to grow priority (GP).

In both the overall and high empathy (HE) cases, the OS condition approached significance in differing from video (V), and V did not statistically differ from the story condition (S) in either scenario. This suggests that the difference in GP was not due to watching a video alone. The S condition showed only 2 GP responses out of 23 participants. While people do report empathizing in the S category, their priorities are not likely to change from this reported empathy. S appeared to be an ineffective way of increasing priority, despite it being by far the most popular way for design research insights to be shared. If participants report LE, we see a similar challenge with video alone, with only 1 out of 11 participants in the V(LE) group and in GP. This suggests that video is most effective when people are able to empathically respond to the video. It may be the case that skin-conductance is particularly effective for helping people who do not empathize easily see the logical importance of improving the customer experience.

When participants report a high empathic response, DS and V do not significantly differ, suggesting the video itself may be enough to grow priority. However, a large effect occurs in OS with 75% of participants growing the priority of the initiative. We can conclude this difference is not because of the skin conductance or the video alone as the effect is larger than V and DS. As this is a post-hoc analysis, future research will be needed to better understand why this OS interpretation had such a strong influence. It may be that OS requires employees to arrive at conclusions on their own, and this sense of ownership helps increase priority. This effect would be similar to the IKEA-Effect - people who build their own IKEA box value it more (Norton et al., 2011). Alternatively, SCR overlay may be effective at targeting cognitive empathy, while we measured affective empathy in the survey. If this is the case, then the OC (HE) condition included people with high emotional empathy and cognitive empathy, which researchers suggest is the most impactful form of empathy (Kerem et al., 2001). Future research should measure people's ability to empathize before varying conditions are introduced to better ascertain how psychophysiology changes empathic response.

An additional challenge to analysis is that empathy scores were measured after each condition. An additional unmeasured factor was people's ability to empathize as a personality trait (e.g. Davis, 1983). Does the OS condition successfully alter people's empathetic response and priorities or, alternatively, is OS most effective with people with empathic personalities?

These results also suggest that the "over-simplification" of psychophysiological data (Beck, 2010) may not be the only driving factor of increased confidence in psychophysiology. When the complexity of interpreting skin conductance was described in detail (OS), % GP was not significantly different than participants given simplified explanations (DS). In fact, when only considering those in high empathy (HE), the open-interpretation (OS) participants were more likely to grow priority (GP) than participants told skin conductance is stress (DS). From a design research stand point, this difference suggests that protecting clients from the "challenges of interpretation" may not only be unnecessary but harmful in generating support.

While this was only the first study on how skin conductance can alter employee's business priorities, the initial results suggest that psychophysiology does have an effect on whether employees will grow priority of an initiative. Depictions of psychophysiology can act as a boundary object that diverse stakeholders can react to. Design researchers should continue to research new ways to generate understandable and motivating data that employees can embrace. Future work should look at what factors may be at play in openly interpreted skin conductance (OS) having such a large effect on people's priorities. What helped LEGO employees embrace a new initiative: the use of technology, the mixing of numbers and story, psychophysiological data, the combination of cognitive and affective empathy, additional factors, or a combination? Regardless, one can conclude that the design initiative itself may not be enough to energize an organization, and designers should look past short stories at new ways to effectively share insights.

Chapter 7: Conclusions

Throughout this dissertation, I showed how thick psychophysiology can be a valuable tool to conduct exploratory psychophysiology. With wireless and wearable sensors becoming the norm, people will now be able to measure psychophysiology in meaningful, open, real-world settings. A natural step will be for psychophysiologists to start asking more open-ended, explorative questions as we try to better understand the mechanisms of emotion and its interactions with the environment.

However, as we begin to conduct research "in the wild," specificity is a serious challenge. I remember starting out, measuring the skin conductance of children with sensory challenges during therapy and telling my professors, "The child is clearly stressed based on her skin conductance." Dr. Picard had to continuously remind me that I did not know it was stress based on skin conductance alone. I hope that others who conduct thick psychophysiology can learn from my initial mistakes and be always questioning the meaning of physiological changes, taking a skeptical stance from the beginning. I outlined a variety of ways inference can occur in ambulatory settings, with thick psychophysiology being a new direction of interpreting explorative, psychophysiological data. Perhaps in a twist of fate, those moments I wanted to call stressful in my initial therapy work, I can now call stressful by fully acknowledging the contextual information, not because of the psychophysiology responses alone, but because of the psychophysiology and the child saying "No, no, no." and the therapist attempting to calm the child down.

I described how one might conduct thick psychophysiology, detailing the challenges and decisions one has to make in combining psychophysiology and observational methods together. Observations can be used to understand both the internal and external contexts of an experience and help provide specificity. Of course, thick psychophysiology is not flawless. I described how movements or external influences can make interpretation of the data difficult. I ended the thick psychophysiology section discussing how one communicates findings to lay audiences. Skin conductance can be a powerful influence, but simultaneously, how can design researchers insure the data is being interpreted correctly? While I emphasize with the pitfalls of thick psychophysiology, I hope this section is also a motivator for researchers who want to understand how people emotionally respond in uncontrolled settings.

The LEGO case study served as a way of illustrating how thick psychophysiology can help design researchers. While I have conducted large psychophysiology studies with Lowes, Best Buy, Hasbro, and Google X, the LEGO Group graciously let me document the process with minimum censorship. The LEGO project had the benefit of being the latest iteration of my practice, after a year of training at IDEO and a year of sociological classes before that. The LEGO case study shows how far thick psychophysiology has developed. Compare the deep detail of Ryan's conversation with his father about having to go back in the instructions to the initial analysis in my master's thesis where I summarize an entire part of therapy as "using the ball pit." The results suggested thick psychophysiology can be fruitful, helping researchers learn about possible nuances about emotional experiences that would have been hard to observe without the lens of psychophysiology.

The New World Symphony case study helped show another way that thick psychophysiology can help with design research. Like the LEGO study, I showed examples of times when psychophysiology helped us identify psychological responses that would be much harder to identify without skin conductance. The New World Symphony also provided an interesting backdrop for a case study as it was a near-ideal location for ambulatory research: stimuli were relatively controlled, participants sat relatively still, and multiple subjects were given the exact same stimuli for easier group comparison. I look forward to the

New World Symphony research growing into a larger project where we can look at how larger groups of people react to the symphony and how the observed effects generalize.

I end this dissertation with a glimpse of future work I hope to do in the form of boundary objects. Pragmatic boundary objects are needed for us to share knowledge across groups. From my personal experience, I knew psychophysiology changed how employees look at customer experience. I have presented to multiple stakeholders who believed focus groups were broken and there was no known way they could understand their customers again. After an hour presentation about skin conductance, their attitudes shifted – perhaps customer experience could be measured after all. In surveys after presentations, employees consistently asked to learn more about skin conductance.

However, these qualitative insights are not sound evidence for a large portion of the audience I care about – employees who are wary of qualitative work. Hence, a quantitative study helped show the effect thick psychophysiology was having on employees across industries. Talking with Hugh Dubberly about my survey results, we concluded that design is ultimately a political process and how a design idea is embraced is a vital part of its success. I often refer to a research project where I learned shoppers were jumping up to see the product display. I spent 5 weeks with additional research, workshops, brainstorms, and prototypes to convince the company that customers jumping to see their product was poor design. I am still unsure if my suggestion of lowering the height of the product display will be executed. In some ways, the execution of design is the more difficult task that hardly is referenced. As the survey results showed, psychophysiology might be a better way to introduce employees to rigorous ethnographic research and help them understand the importance of customer pain points.

7.1 FUTURE WORK

This dissertation leaves more unanswered questions than answers, which I believe is a good outcome of explorative research. For me personally, I am starting a consulting company directed at helping organizations better empathize and understand their customers, and many of my future research questions follow that line of thinking.

Question 1: What makes effective boundary objects for design? We showed that depictions of psychophysiology signals can have a strong effect with up to 75% of empathic viewers increasing the priority of an initiative when skin conductance was shown. Why was such a large effect present? From my work at IDEO I know that co-creating with clients, having immersive environments, and telling stories are effective ways of communicating design insights, but how effective are they? And what makes them effective? How can skin conductance be even more effective as a boundary object for those less inclined to consider customer experience?

I envision a follow up study to the survey from Chapter 6 where we measure people's ability to empathize as a personality trait first, measure the effect in a different setting, and measure cognitive empathy as well as emotional empathy.

Question 2: I would like to see a full ethnography be done with skin conductance sensors as a lens. Perhaps an anthropologist could look at the day to day interactions of children in a classroom and ask "How do they regulate their emotions?" How do drivers curb their road rage? What are the moments of stress for single mothers working part time jobs? Why are people so excited by roller coasters? The number of emotional questions to be explored is vast, and I hope to see some ethnographers embrace thick psychophysiology as a tool and begin to explore emotions in new ways.

Question 3: What are common emotional patterns in service design? I have started working with a handful of service industry companies, and I find that the outcomes in the different settings have similar forms. If participants are stressed early on, they are likely to be more stressed by future events. Too many choices consistently overwhelm shoppers. Interacting with products is a great way to involve customers, but companies often fall flat on execution. Interfacing between the web and reality in most retail spaces is broken. As the number of clients I work with increases, I hope to start documenting reoccurring patterns so findings can be generalized and shared.

Question 4: What does psychophysiology help researchers discover that they would have otherwise missed? I know skin conductance has helped me arrive at many of the insights that I list throughout the dissertation. I wonder if I would have come to the same conclusions without the aid of skin conductance. This type of study lends itself well to a controlled test, ideally in an emotional environment. I have been hoping to measure the stress of parents shopping at Toys R' Us or the stress of talking about finances at Fidelity and seeing what skin conductance not only helps communicate, but also discover. Compared to observations and interviews, I expect psychophysiology helps identify some types of emotional responses better.

Question 5: How could therapy be redesigned to be more human centered? When I first started working with occupational therapists, I only had a glimpse of the power of design. Now, with a more applied toolset, I would like to see if we could prototype and redesign the therapies for children, helping therapists not only understand the emotional reactions of patients, but build out new ways of addressing those emotional experiences.

7.2 CONCLUDING REMARKS

I end every story with my father Dana. My dad is a school bus driver in Summit County, Colorado where there tends to be snow on the ground most of the year. Every time I call my father, he updates me on his passengers. Jimmy got in a fight with Sam today or Susie did really well on her math test. My father cares about his customers. On the last day of school, he brought a box of pinwheels and gave one to each child. As they drove down the highway, the children put the pinwheels out the window and started smiling and laughing as my Dad said good bye for the summer. I want the Best Buys and Comcasts of the world to care and understand their customers in the same way my Dad understands his. I believe thick psychophysiology can help bridge decision makers with the customers they serve. Today, there is not enough empathy in the world and far too few pinwheels; let’s change that.

Bibliography

Adam, E. K., & Gunnar, M. R. (2001). Relationship functioning and home and work demands predict individual differences in diurnal cortisol patterns in women. *Psychoneuroendocrinology, 26*(2), 189-208.

Anderson, S. F., & Lawler, K. A. (1995). The anger recall interview and cardiovascular reactivity in women: An examination of context and experience. *Journal of Psychosomatic Research, 39*(3), 335-343.

Antrobus, J., Coleman, R., & Singer, J. (1967). Signal-detection performance by subjects differing in predisposition to daydreaming. *Journal of Consulting Psychology, 31*(5), 487.

Arrow, K. J. (1994). Methodological individualism and social knowledge. *American Economic Review, 84*(2), 1-9.

Babbie, E. R. (2013). *The practice of social research* Cengage Learning.

Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality, 55*(1), 19-39.

Bearden, S., & Moffatt, R. (2001). V̇ o 2 and heart rate kinetics in cycling: Transitions from an elevated baseline. *Journal of Applied Physiology, 90*(6), 2081-2087.

Bechky, B. A. (2003). Sharing meaning across occupational communities: The transformation of understanding on a production floor. *Organization Science, 14*(3), 312-330.

Beck, D. M. (2010). The appeal of the brain in the popular press. *Perspectives on Psychological Science, 5*(6), 762-766.

Ben-Shakhar, G. (1985). Standardization within individuals: A simple method to neutralize individual differences in skin conductance. *Psychophysiology, 22*(3), 292-299.

Blix, A. S., Stromme, S. B., & Ursin, H. (1974). Additional heart rate--an indicator of psychological activation. *Aerospace Medicine, 45*(11), 1219-1222.

Boucsein, W. (2011). *Electrodermal activity* Springer Verlag.

Boucsein, W., Haarmann, A., & Schaefer, F. (2007). Combining skin conductance and heart rate variability for adaptive automation during simulated IFR flight. *Engineering Psychology and Cognitive Ergonomics* 639-647.

Boucsein, W., Schaefer, F., & Sommer, T. (2001). Electrodermal long-term monitoring in everyday life. *Progress in Ambulatory Assessment (S.549-560).Göttingen: Hogrefe,*

Briggs, J. L. (1970). *Never in anger: Portrait of an eskimo family* Harvard University Press.

Bueno, M., & Podolsky, J. (2010). Experiencing services before they become real. touchpoint. *The Journal of Service Design, 2*(2)

Cacioppo, J. T., & Tassinary, L. G. (1990). *Principles of psychophysiology: Physical, social, and inferential elements.* Cambridge University Press.

Cacioppo, J. T., Martzke, J. S., Petty, R. E., & Tassinary, L. G. (1988). Specific forms of facial EMG response index emotions during an interview: From darwin to the continuous flow hypothesis of affect-laden information processing. *Journal of Personality and Social Psychology, 54*(4), 592.

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *The Journal of Political Economy* 1232-1254.

Carlile, P. R. (2002). A pragmatic view of knowledge and boundaries: Boundary objects in new product development. *Organization Science, 13*(4), 442-455.

Carlile, P. R., & Rebentisch, E. S. (2003). Into the black box: The knowledge transformation cycle. *Management Science, 49*(9), 1180-1195.

Christie, I. C., & Friedman, B. H. (2004). Autonomic specificity of discrete emotion and dimensions of affective space: A multivariate approach. *International Journal of Psychophysiology, 51*(2), 143-153.

Cohen, D., Nisbett, R. E., Bowdle, B. F., & Schwarz, N. (1996). Insult, aggression, and the southern culture of honor: An" experimental ethnography.". *Journal of Personality and Social Psychology, 70*(5), 945.

Critchley, H. D. (2002a). Electrodermal responses: What happens in the brain. .*The Neuroscientist, 8*, 132–142.

Critchley, H. D. (2002b). Electrodermal responses: What happens in the brain. *The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry, 8*(2), 132-142.

Daily, S. B. (2010). *More than a Feeling: Technology-Infused Learning Environments to Support the Development of Empathy,*

Daily, S. B., & Picard, R. (2004). INNER-active journal. *Proceedings of the 1st ACM Workshop on Story Representation, Mechanism and Context*, 51-54.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113.

Debener, S., Minow, F., Emkes, R., Gandras, K., & Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology, 49*(11), 1617-1621.

Diefenbach, T. (2009). Are case studies more than sophisticated storytelling?: Methodological problems of qualitative empirical research mainly based on semi-structured interviews. *Quality & Quantity, 43*(6), 875-894.

Dozier, M., & Kobak, R. R. (1992). Psychophysiology in attachment interviews: Converging evidence for deactivating strategies. *Child Development, 63*(6), 1473-1480.

Dubberly, H., & Evenson, S. (2008). On modeling the analysis-systhesis bridge model. *Interactions, 15*(2), 57-61.

Dubberly, H. (2012). What can steve jobs and jonathan ive teach us about designing? *Interactions, 19*(3), 82-85.

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin, 101*(1), 91.

Fahrenberg, J. (2006). Assessment in daily life. A review of computer-assisted methodologies and applications in psychology and psychophysiology, years 2000-2005.

Fahrenberg, J. (1996). Ambulatory assessment: Issues and perspectives
. *Ambulatory Assessment: Computer-Assisted Psychological and Psychophysiological Methods in Monitoring and Field Studies* 3-20.

Feshbach, N. D. (1975). Empathy in children: Some theoretical and empirical considerations. *The Counseling Psychologist,*

Fletcher, R., Dobson, K., Goodwin, M., Eydgahi, H., Wilder-Smith, O., Fernholz, D., . . . Picard, R. (2010). iCalm: Wearable sensor and network architecture for wirelessly communicating and logging autonomic activity. *Information Technology in Biomedicine, IEEE Transactions On, 14*(2), 215-223.

Foerster, F., & Fahrenberg, J. (2000). Motion pattern and posture: Correctly assessed by calibrated accelerometers. *Behavior Research Methods, Instruments, & Computers, 32*(3), 450-457.

Freeze, K. J., & Chung, K. (2008). Design strategy at samsung electronics: Becoming a top-tier company. *Case Study, Design Management Institute* 441.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. *Culture: Critical Concepts in Sociology* 173-196.

Gladstein, G. A. (1983). Understanding empathy: Integrating counseling, developmental, and social psychology perspectives. *Journal of Counseling Psychology, 30*(4), 467.

Glaser, B. G., & Strauss, A. L. (2009). *The discovery of grounded theory: Strategies for qualitative research* Transaction Publishers.

Goodfellow, T. (2012). ***Three moves ahead episode 185 – class in session***

Gould, R., & Sigall, H. (1977). The effects of empathy and outcome on attribution: An examination of the divergent-perspectives hypothesis. *Journal of Experimental Social Psychology, 13*(5), 480-491.

Gray, M. A., Harrison, N. A., Wiens, S., & Critchley, H. D. (2007). Modulation of emotional appraisal by false physiological feedback during fMRI. *PLoS One, 2*(6), e546.

Greco, M., & Baenninger, R. (1991). Effects of yawning and related activities on skin conductance and heart rate. *Physiology & Behavior, 50*(5), 1067-1069.

Green, W. S., & Jordan, P. W. (2002). *Pleasure with products: Beyond usability* CRC.

Groeppel-Klein, A. (2005). Arousal and consumer in-store behavior. *Brain Research Bulletin, 67*(5), 428-437.

Hansson, G., Asterland, P., Holmer, N., & Skerfving, S. (2001). Validity and reliability of triaxial accelerometers for inclinometry in posture analysis. *Medical and Biological Engineering and Computing, 39*(4), 405-413.

Harris, V. A., & Jellison, J. M. (1971). Fear-arousing communications, false physiological feedback, and the acceptance of recommendations. *Journal of Experimental Social Psychology, 7*(3), 269-279.

Hazlett, R. L. (2006). Measuring emotional valence during interactive experiences: Boys at video game play. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 1023-1026.

Hazlett, R. L., & Hazlett, S. Y. (1999). Emotional response to television commercials: Facial EMG vs. self-report. *Journal of Advertising Research, 39,* 7-24.

Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *Intelligent Transportation Systems, IEEE Transactions On, 6*(2), 156-166.

Hedman, E. (2010). *In-situ measurement of electrodermal activity during occupational therapy.* (Unpublished Masters). Massachusetts Institute of Technology,

Hedman, E. (2011). The frustration of learning monopoly: The emotional tension of entering a new game encounter. *The Ethnographic Praxis in Industry Conference, September 18-21,* Boulder.

Hedman, E. B., Quigley, K. S., & Picard, R. W. (2012). Using ranked scrs in ambulatory measurements: a new approach to distinguishing real-world salient events. *PSYCHOPHYSIOLOGY 49* S118-S118.

Hedman, E., Miller, L. u., Schoen, S., Nielsen, D., Goodwin, M., & Picard, R. (2012). Measuring autonomic arousal during therapy. *Proceedings of 8th International Design and Emotion Conference*, London.

Hirsch, J., & Bishop, B. (1981). Respiratory sinus arrhythmia in humans: How breathing pattern modulates heart rate. *American Journal of Physiology-Heart and Circulatory Physiology, 241*(4), H620-H629.

Hjortskov, N., Garde, A. H., Ørbæk, P., & Hansen, Å M. (2004). Evaluation of salivary cortisol as a biomarker of self-reported mental stress in field studies. *Stress and Health, 20*(2), 91-98.

Hochschild, A. R. (2003). *The managed heart: Commercialization of human feeling* Univ of California Pr.

Hot, P., Naveteur, J., Leconte, P., & Sequeira, H. (1999). Diurnal variations of tonic electrodermal activity. *International Journal of Psychophysiology, 33*(3), 223-230.

James, J. A., & Daly, M. d. B. (1969). Nasal reflexes. *Proceedings of the Royal Society of Medicine, 62*(12), 1287.

Jones, E. E., & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin, 76*(5), 349.

Kamarck, T. W., Schwartz, J. E., Janicki, D. L., Shiffman, S., & Raynor, D. A. (2003). Correspondence between laboratory and ambulatory measures of cardiovascular reactivity: A multilevel modeling approach. *Psychophysiology, 40*(5), 675-683.

Kanning, M. (2012). Using objective, real-time measures to investigate the effect of actual physical activity on affective states in everyday life differentiating the contexts of working and leisure time in a sample with students. *Frontiers in Psychology, 3*

Katz, J. (2001). *How emotions work* University of Chicago Press.

Katz, R. L. (1963). *Empathy: Its nature and uses* Free Press of Glencoe London.

Kerem, E., Fishman, N., & Josselson, R. (2001). The experience of empathy in everyday relationships: Cognitive and affective elements. *Journal of Social and Personal Relationships, 18*(5), 709-729.

Kolodyazhniy, V., Kreibig, S. D., Gross, J. J., Roth, W. T., & Wilhelm, F. H. (2011). An affective computing approach to physiological emotion specificity: Toward subject-independent and stimulus-independent classification of film-induced emotions. *Psychophysiology, 48*(7), 908-922.

Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. *Multiple-Task Performance* 279-328.

Krebs, D. (1975). Empathy and altruism. *Journal of Personality and Social Psychology, 32*(6), 1134.

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., & Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear-and sadness-inducing films. *Psychophysiology, 44*(5), 787-806.

LaBar, K. S., & Phelps, E. A. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science, 9*(6), 490-493.

LaBarbera, P., & Tucciarone, J. (1995). GSR reconsidered: A behavior-based approach to evaluating and improving the sales potency of advertising. *Journal of Advertising Research,*

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261-273.

Leavitt, J. (1996). Meaning and feeling in the anthropology of emotions. *American Ethnologist, 23*(3), 514-539.

Lee, M. H., & Nicholls, H. R. (1999). Review article tactile sensing for mechatronics—a state of the art survey. *Mechatronics, 9*(1), 1-31.

Leonard, D., & Rayport, J. F. (1997). Spark innovation through empathic design. *Harvard Business Review, 75*, 102-115.

Levine, L. J., Lench, H. C., & Safer, M. A. (2009). Functions of remembering and misremembering emotion. *Applied Cognitive Psychology, 23*(8), 1059-1075.

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition, 22*(5: Special issue), 530-554.

Light, L. (1993). At the center of it all is the brand. *Advertising Age, 64*(13), 22.

Lorussi, F., Rocchia, W., Scilingo, E. P., Tognetti, A., & De Rossi, D. (2004). Wearable, redundant fabric-based sensor arrays for reconstruction of body segment posture. *Sensors Journal, IEEE, 4*(6), 807-818.

Lutz, C., & White, G. M. (1986). The anthropology of emotions. *Annual Review of Anthropology, 15*, 405-436.

Lutz, C. A., & Abu-Lughod, L. E. (1990). Language and the politics of emotion. *This Book Grew Out of a Session at the 1987 Annual Meeting of the American Anthropological Association Called" Emotion and Discourse.",*

Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology, 8*(5), 656-672.

Lynn, R. (1966). Attention, arousal and the orientation reaction.

Mallion, J., Baguet, J., Siché, J., Tremel, F., & De Gaudemaris, R. (1999). Clinical value of ambulatory blood pressure monitoring. *Journal of Hypertension, 17*(5), 585-595.

Mangina, C. A., & Beuzeron-Mangina, J. H. (1996). Direct electrical stimulation of specific human brain structures and bilateral electrodermal activity. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology, 22*(1-2), 1-8.

Martins, L. L., & Kambil, A. (1999). Research notes: Looking back and thinking ahead: Effects of prior success on managers' interpretations of new information technologies. *Academy of Management Journal, 42*(6), 652-661.

McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107*(1), 343-352.

McCabe, D. P., Castel, A. D., & Rhodes, M. G. (2011). The influence of fMRI lie detection evidence on juror Decision-Making. *Behavioral Sciences & the Law, 29*(4), 566-577.

Meng, C., LI, M., & LI, A. (2004). Digital temperature and humidity sensor SHT11 based on I~ 2C bus and its application in the single-chip microcomputer system [J]. *International Electronic Elements, 2*, 016.

Mirza-Babaei, P., Long, S., Foley, E., & McAllister, G. (2011). Understanding the contribution of biometrics to games user research. *Proc. DIGRA,*

Morris, A. L., Cleary, A. M., & Still, M. L. (2008). The role of autonomic arousal in feelings of familiarity. *Consciousness and Cognition, 17*(4), 1378-1385.

Myrtek, M. (2004). *Heart and emotion: Ambulatory monitoring studies in everyday life.* Hogrefe & Huber Publishers.

North, A. C., & Hargreaves, D. J. (1997). Experimental aesthetics and everyday music listening.

Norton, M., Mochon, D., & Ariely, D. (2011). The'IKEA effect': When labor leads to love. *Harvard Business School Marketing Unit Working Paper,* (11-091)

Pfaltz, M. C., Michael, T., Grossman, P., Blechert, J., & Wilhelm, F. H. (2009). Respiratory pathophysiology of panic disorder: An ambulatory monitoring study. *Psychosomatic Medicine, 71*(8), 869-876.

Picard, R. W., Fedor, S., & Ayzenberg, Y. (Manuscript submitted for publication.). Multiple arousal theory and daily-life electrodermal activity asymmetry
. *Emotion Review.,*

Pickering, T. G., Eguchi, K., & Kario, K. (2007). Masked hypertension: A review. *Hypertension Research, 30*(6), 479.

Poh, M. Z., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *Biomedical Engineering, IEEE Transactions On, 57*(5), 1243-1252.

Pomeranz, B., Macaulay, R., Caudill, M. A., Kutz, I., Adam, D., Gordon, D., . . . Cohen, R. J. (1985). Assessment of autonomic function in humans by heart rate spectral analysis. *American Journal of Physiology-Heart and Circulatory Physiology, 248*(1), H151-H153.

Quigley, K. S., & Barrett, L. F. (2014). Is there consistency and specificity of autonomic changes during emotional episodes? guidance from the conceptual act theory and psychophysiology. *Biological Psychology,*

Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S., & Keltikangas-Järvinen, L. (2008). The psychophysiology of james bond: Phasic emotional responses to violent video game events. *Emotion, 8*(1), 114.

Regan, D. T., & Totten, J. (1975). Empathy and attribution: Turning observers into actors. *Journal of Personality and Social Psychology, 32*(5), 850.

Rosaldo, M. Z. (1983). The shame of headhunters and the autonomy of self. *Ethos, 11*(3), 135-151.

Schafer, R. (1959). Generative empathy in the treatment situation. *The Psychoanalytic Quarterly,*

Scheirer, J., Fernandez, R., Klein, J., & Picard, R. W. (2002). Frustrating the user on purpose: A step toward building an affective computer. *Interacting with Computers, 14*(2), 93-118.

Seemann, J. (2012). Hybrid insights: Where the quantitative meets the qualitative. *Rotman Magazine, Iss* 57-61.

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain : A Journal of Neurology, 132*(Pt 3), 617-627. doi:10.1093/brain/awn279; 10.1093/brain/awn279

Sigall, H., & Page, R. (1971). Current stereotypes: A little fading, a little faking. *Journal of Personality and Social Psychology, 18*(2), 247.

Singer, J. L., Greenberg, S., & Antrobus, J. S. (1971). Looking with the mind's eye: Experimental studies of ocular motility during daydreaming and mental arithmetic. *Transactions of the New York Academy of Sciences, 33*(7 Series II), 694-709.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946.

Spitzer, S. B., Llabre, M. M., Ironson, G. H., Gellman, M. D., & Schneiderman, N. (1992). The influence of social situations on ambulatory blood pressure. *Psychosomatic Medicine, 54*(1), 79-86.

Springer, J., Müller, T., LANGNERT, L. H., & Beitz, W. (1989). Stress and strain caused by CAD-work. results of a laboratory study. *BERLINGUET L, Berthelette D.(Éds)-Work with Display Units, 89*, 231-238.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology,translations' and boundary objects: Amateurs and professionals in berkeley's museum of vertebrate zoology, 1907-39. *Social Studies of Science, 19*(3), 387-420.

Stephens, C. L., Christie, I. C., & Friedman, B. H. (2010). Autonomic specificity of basic emotions: Evidence from pattern classification and cluster analysis. *Biological Psychology, 84*(3), 463-473.

Strayer, J. (1987). Affective and cognitive perspectives on empathy.

Stuiver, A., & Mulder, B. (2009). Artefact-free real-time computation of cardiovascular measures. *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference On,* 1-6.

Thayer, R. (1989). The biopsychology of mood and activation.

Thornton, J. A. (2005). Fear patterns in anti-speeding advertisements: Effects on simulated driver-behaviour.

Toi, M., & Batson, C. D. (1982). More evidence that empathy is a source of altruistic motivation. *Journal of Personality and Social Psychology, 43*(2), 281.

Trost, S. G., McIver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and Science in Sports and Exercise, 37*(11), S531.

Turpin, G. (1986). Effects of stimulus intensity on autonomic responding: The problem of differentiating orienting and defense reflexes. *Psychophysiology, 23*(1), 1-14.

van Doornen, L. J., Houtveen, J. H., Langelaan, S., Bakker, A. B., van Rhenen, W., & Schaufeli, W. B. (2009). Burnout versus work engagement in their effects on 24-hour ambulatory monitored cardiac autonomic function. *Stress and Health, 25*(4), 323-331.

Waters, W. F., Koresko, R. L., Rossie, G. V., & Hackley, S. A. (1979). Short-, medium-, and Long-Term relationships among meteorological and electrodermal variables. *Psychophysiology, 16*(5), 445-451.

Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience, 20*(3), 470-477.

Westenberg, P. M., Bokhorst, C. L., Miers, A. C., Sumter, S. R., Kallen, V. L., van Pelt, J., & Blöte, A. W. (2009). A prepared speech in front of a pre-recorded audience: Subjective, physiological, and neuroendocrine responses to the leiden public speaking task. *Biological Psychology, 82*(2), 116-124.

Wild, J., Clark, D. M., Ehlers, A., & McManus, F. (2008). Perception of arousal in social anxiety: Effects of false feedback during a social interaction. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(2), 102-116.

Wilhelm, F. H., & Grossman, P. (2010). Emotions beyond the laboratory: Theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment. *Biological Psychology, 84*(3), 552-569.

Wilhelm, F. H., & Roth, W. T. (1998). Taking the laboratory to the skies: Ambulatory assessment of self-report, autonomic, and respiratory responses in flying phobia. *Psychophysiology, 35*(5), 596-606.

Wilhelm, F. H., & Roth, W. T. (1996). Ambulatory assessment of clinical anxiety. *Ambulatory Assessment: Computerassisted Psychological and Psychophysiological Methods in Monitoring and Field Studies* 317-345.

Wilhelm, F. H., Trabert, W., & Roth, W. T. (2001). Physiologic instability in panic disorder and generalized anxiety disorder. *Biological Psychiatry, 49*(7), 596-605.

Wilson, G. (2001). In-flight psychophysiological monitoring. *Progress in Ambulatory Monitoring* 435-454.

Wulff, H. (2007). *Introduction: The cultural study of mood and meaning* Berg Publishers, Oxford.

Yamasaki, Y., Kodama, M., Matsuhisa, M., Kishimoto, M., Ozaki, H., Tani, A., . . . Kamada, T. (1996). Diurnal heart rate variability in healthy subjects: Effects of aging and sex difference. *American Journal of Physiology-Heart and Circulatory Physiology, 271*(1), H303-H310.

Zeier, H., Häseli, A., & Fischer, J. (2001). Heart rate monitoring in an academic test situation. *Progress in Ambulatory Assessment* 387-398.

Appendices

APPENDIX A: LEGO BOUNDARY OBJECT SURVEY

This survey is voluntary and you can choose to not answer any questions or exit the survey with no consequence. Additionally, all your answers will be anonymous.

Initiative 1:

When mothers help build, children are less stressed and can build more challenging LEGO sets. Unfortunately, mothers rarely build with their children. We should develop new ways to better include mothers and improve the building experience.

1. How high of a priority should "Creating a Mom-Friendly LEGO Building Experience" be for the LEGO Group?

	1	2	3	4	5	
Lowest Priority	○	○	○	○	○	Highest Priority

Initiative 2:

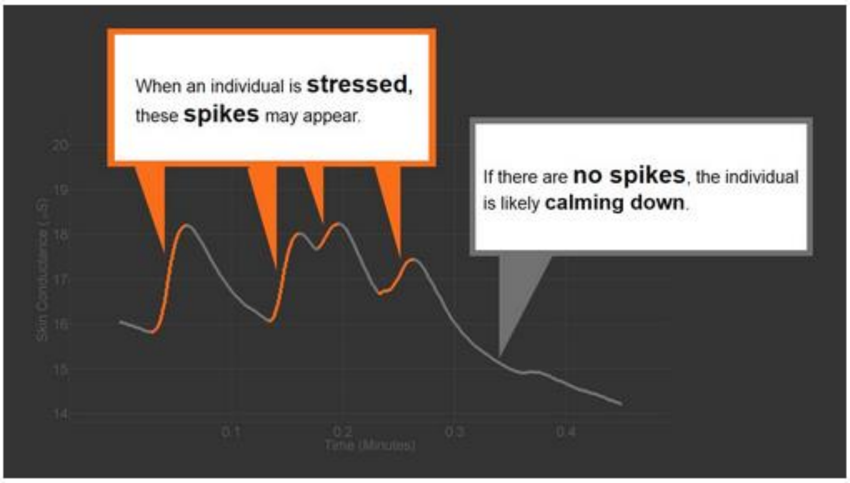
Researching the best ways to market to children on the tablet could give the LEGO group an edge in communicating its value and correspondingly increase its sales.

2. How high of a priority should "Improving Marketing on Tablets" be for the LEGO Group?

	1	2	3	4	5	
Lowest Priority	○	○	○	○	○	Highest Priority

In some of the initiatives presented, children wore a new technology, the MOXO sensor. Please familiarize yourself with the MOXO sensor below.

THE MOXO sensor measures children's emotional reactions from their skin conductance, a marker of the sympathetic nervous system. When an orange spike occurs, the child is more likely to be stressed. If there are no orange lines, the child is likely calming down. In the image below a child becomes stressed then relaxed.



The MOXO Sensor

Initiative 1: Creating a Mom-Friendly LEGO Building Experience

When mothers help build, children are often less stressed and can build more challenging LEGO sets. Unfortunately, mothers rarely build with their children. We should develop new ways to better include mothers and improve the building experience.

So far, we interviewed 10 families about playing with LEGO bricks. Take for example, Matthew, a 6 year old boy from Vejle.

When Matthew's grandparents gave him his first Technic, his mother sat alongside him and supported him at the challenging parts. Without her, Matthew would have likely become overwhelmed, and may not have finished.

3. In three words, how would you describe your reaction to Matthew's story?

Word 1

Word 2

Word 3

4. Now, how high of a priority should "Creating a Mom-Friendly LEGO Building Experience" be for the LEGO Group?

	1	2	3	4	5	
Not a Priority	○	○	○	○	○	Top Priority

5. How stressed did you think Matthew was when he built Technic without his mother?

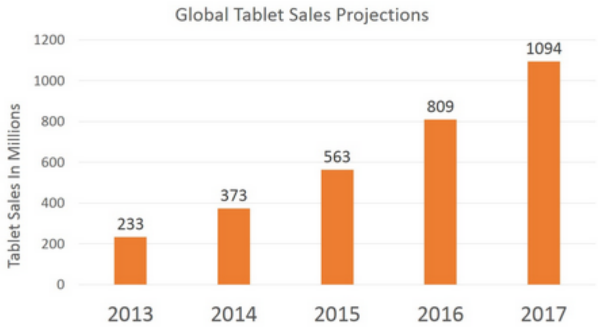
	1	2	3	4	5	
Not at All	○	○	○	○	○	Extremely

6. How much did you experience each of these emotions when learning about Matthew?

	Not at All 1	2	3	4	5	6	Extremely 7
Compassionate	○	○	○	○	○	○	○
Concerned	○	○	○	○	○	○	○
Tender	○	○	○	○	○	○	○
Inspired	○	○	○	○	○	○	○
Softhearted	○	○	○	○	○	○	○
Sad	○	○	○	○	○	○	○
Sympathetic	○	○	○	○	○	○	○
Interested	○	○	○	○	○	○	○
Moved	○	○	○	○	○	○	○
Confused	○	○	○	○	○	○	○
Warm	○	○	○	○	○	○	○

Initiative 2: Tablet Marketing

Kids are spending more and more time on tablets. In 2013, children who play with tablets daily were estimated to spend 7 hours a week on those tablets. It is important for LEGO marketing to meet children where they are, and for the near future that is on the tablet. As tablets have a touch element, children can now interact with ads in new ways, and marketing can be more closely integrated in apps. Researching the best ways to market to children on the tablet could give the LEGO group an edge in communicating its value and correspondingly increase its sales.



7. Which initiative should we do? This initiative could be funded up to 2,500,000 DKK. Additionally, up to a full day workshop, specifically for your department, may be held on this topic. (Your answer will be anonymous and you can choose either project.)

- Initiative 1: Creating a Mom-Friendly LEGO Building Experience
- Initiative 2: Tablet Marketing

8. How confident are you that you selected the most important initiative?

	1	2	3	4	5	
Not at All	○	○	○	○	○	**Extremely**

9. If the "Mom Friendly LEGO Building" Initiative is chosen, how many hours of the department wide workshop would you like to attend?

- 0 hours
- 1 hour
- Half Day
- Full Day

Background Questions

10. Would you like to receive emails about developments in

- ☐ Mom Friendly LEGO Building
- ☐ Tablet Marketing

11. We may be looking for after-work volunteers for the "Creating a Mom-Friendly LEGO Building Experience" initiative. Would you like to sign up for the volunteer mailing list?

- ○ Yes
- ○ No

12. If you answered yes, how many hours do you think you might be able to volunteer each week?

13. Gender?

- ○ Male
- ○ Female

14. Age?

- ○ under 18
- ○ 18-24
- ○ 25-34
- ○ 35-54
- ○ 55+

16. Do you have children with the following ages?

- ☐ Boy(s) Under 5
- ☐ Boy(s) 5-12
- ☐ Boy(s) Older than 12
- ☐ Girl(s) Under 5
- ☐ Girl(s) 5-12
- ☐ Girl(s) Older than 12

17. What best describes your position at LEGO?

- ○ Support
- ○ Professional
- ○ Lead Professional
- ○ Manager
- ○ Senior Manager
- ○ Director
- ○ Senior Director
- ○ Vice President
- ○ Senior Vice President
- ○ Executive Vice President

18. Have you been to a talk by Elliott Hedman from MIT about these MOXO sensors?

- ○ Yes
- ○ No

19. What LEGO set was Matthew playing with?

- Mindstorms
- Chima
- Technic
- I don't remember

---

20. Among daily users, how many hours each week do children play with tablets?

- 3 hour
- 7 hours
- 10 hours
- I don't remember

---

21. If you indicated you would like us to contact you, please provide your email below

21. Survey Complete!

Please keep in mind that the presented projects were not real and do not directly reflect what the LEGO Group or the MIT Media Lab plans on doing in the future. Rather, we wanted your opinion on what kind of future projects you would support and how you made that decision.

Please do not share anything about this survey with your colleagues until after this week. We want to make sure we get employee's initial thoughts.

What would you like to do now?

- Learn More about the Benefits of Including Mom in Play (this will take you to a new page)
- Exit

Initiative 1: Creating a Mom-Friendly LEGO Building Experience

When mothers help build, children are often less stressed and can build more challenging LEGO sets. Unfortunately, mothers rarely build with their children. We should develop new ways to better include mothers and improve the building experience.

So far, we videotaped 10 children as they played with LEGO Technic. Take for example this video of Matthew, a 6 year old boy from Vejle. Matthew builds his fist Technic alone, and the instructions become too hard. Soon after the video clip, Matthew stands up and asks his mother if he can quit playing. When his mother helps him, he stays calm.

Please Watch this Short Video.

3. Were you able to:

- ☐ Watch the movie
- ☐ Hear Sound

Initiative 1: Creating a Mom-Friendly LEGO Building Experience

When mothers help build, children are often less stressed and can build more challenging LEGO sets. Unfortunately, mothers rarely build with their children. We should develop new ways to better include mothers and improve the building experience.

So far, we had 10 children wear the MOXO sensors as they played with Technic in their home. Take for example this video of Matthew, a 6 year old boy from Vejle. Matthew builds his fist Technic alone, and the instructions become too hard. Soon after the video clip, Matthew stands up and asks his mother if he can quit playing. When his mother helps him, he stays calm.

Please Watch this Short Video with Matthew's Skin Conductance overlayed.

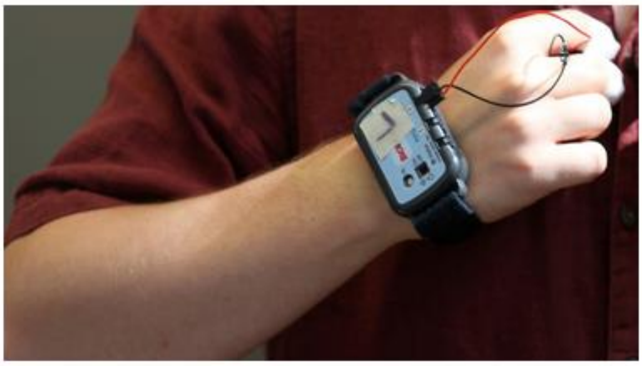
3. Were you able to:

- Watch the movie
- Hear Sound

In some of the initiatives presented, children wore a new technology, the MOXO sensor. Please familiarize yourself with the MOXO sensor below.

THE MOXO sensor measures children's emotional reactions from their skin conductance, a marker of the sympathetic nervous system. When a child becomes stressed, excited, thinks hard, breathes hard, or physically moves, the MOXO data will spike up (orange lines). You will need to look at contextual clues like what the child says and how he behaves to determine why spikes might have occurred. If there are no orange lines, the child is likely calming down.

Skin conductance responses can occur for a variety of reasons.



The MOXO Sensor

APPENDIX B: FULL LOGIT MODEL OF CONDITION ON GP

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	GroupMovie	1.104	.853	1.676	1	.195	3.016
	GroupSCRStress	1.818	.825	4.849	1	.028	6.157
	GroupSCRInterp	2.044	.833	6.018	1	.014	7.724
	Female by Has_boys	-.375	.825	.207	1	.649	.687
	Female	.505	.541	.870	1	.351	1.656
	Has_boys	.075	.592	.016	1	.899	1.078
	Constant	-2.424	.870	7.769	1	.005	.089

a. Variable(s) entered on step 1: GroupMovie, GroupSCRStress, GroupSCRInterp, Female * Has_boys , Female, Has_boys.