THE BRIDGE: LINKING MOOD INDUCTION, SELF-REPORT, AND PSYCHOPHYSIOLOGY TO VOCABULARY LEARNING ON A PAIRED-ASSOCIATES LEARNING TASK

By

Jessica Kate Fox

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Second Language Studies—Doctor of Philosophy

2017

ABSTRACT

THE BRIDGE: LINKING MOOD INDUCTION, SELF-REPORT, AND PSYCHOPHYSIOLOGY TO VOCABULARY LEARNING ON A PAIRED-ASSOCIATES LEARNING TASK

By

Jessica Kate Fox

Researchers in the field of second language acquisition continue to establish links between cognition and emotion (Dewaele, 2013; MacIntyre, 2002; MacIntyre & Gardner, 1989, 1991b, 1994; Segalowitz & Trofimovich, 2011). The purpose of the present study is to investigate to what extent physiological and self-report measures predict vocabulary language learning. This present study is inspired by *hot cognition*, cognitive processing influenced by emotions (Brand, 1987; Pekrun, 2006; Wolfe, 2006). Two groups of thirty-five adult language learners were placed in a negative experimental group or a neutral comparison group and exposed to a series of mood-inducing video-only film clips (Carvalho, Leite, Galdo-Álvarez, & Gonçalves, 2012) after which they learned the forms and meanings of 24 Indonesian concrete nouns. Participant physiological response measures (heart rate, heart rate variability, and skin conductance levels) were collected during baseline and film-viewing periods; additional data collected included periodic emotional self-reports, performance on immediate vocabulary-learning posttests, and a battery of anxiety questionnaires. Findings revealed that changes in heart rate and skin conductance levels influenced performance on the paired-associates vocabulary-learning task. Additionally, the skin conductance measure predicted vocabulary learning when the effects of mood induction and all other known individual differences were controlled for.

Keywords: psychophysiology, hot cognition, vocabulary acquisition

ACKNOWLEDGMENTS

Life is short,

And we do not have much time to gladden the hearts [and enlighten the minds]

of those who make the journey with us.

So… be swift to love, and make haste to be kind (Amiel, 2010).

Life is certainly short, as were the three years of this graduate program. Nonetheless, I would like to thank a number of individuals who traveled the way with me, gladdening my heart and enlightening my mind.

First and foremost, this dissertation would not exist without the direction of Dr. Aline Godfroid, the chair of my dissertation committee, advisor, psycholinguistics professor, and Dutch-speaking friend. Thank you for leading by example in your commitment to rigorous research and sincere care for your students. Next, I owe a special thanks to Dr. Jason Moser, who also made this project possible by generously sharing data collection tools, and brilliant psychology graduate students, and the psychophysiology lab. I am particularly grateful to Dr. Tim Moran and (Dr.) Hans Schroeder; both of these scholars have demonstrated such generosity of their time and knowledge. I would also like to thank the two remaining committee members, Drs. Susan Gass and Paula Winke, for their insight and leadership throughout the course of my doctoral career.

To (Dr.) Zachary Miller, my *hot cognition* psycholinguistics project partner in crime: thank you for your excellent organization, hours of data collection, and get-it-done attitude. Many thanks to Dustin Crowther, Jeffrey Maloney, and Magda Tigchelaar for serving as pilot participants. I would like to thank Tatyana Li and all the members of the MSU Center for

Statistical Training and Consulting (CSTAT) for providing invaluable statistical counseling, unfailing patience, and wizardly skills with the *R* statistical package. Another wizard I would like to thank is Russell Werner, the Random Information Generator, for his never-ending willingness to push the magic button that solved the oft-encountered technical issues. More generally, I would like to thank all of the research project participants involved in the present study, the Second Language Studies program for providing research participant funding, the MSU College of Arts and Letters 2015 Summer Support Research Fellowship, and the Second Language Studies program faculty members and the SLS and MA TESOL student body for your support and camaraderie along the way.

On a more personal note, I would like to thank friends and family who especially gladdened my heart during this leg of the journey: Amber, Beth, Brenda K., Brenda M., Denise, Heather, Heidi, Naomi, Pat, Rachel, Sam, and Stephanie. Thank you all for your hospitality of heart and home. Furthermore, I would like to offer an enormous thank you to the "usual suspects:" Emily, Erin, Eva, Laura, Lauren, Megan, and Spencer. You have been more than just band mates, classmates, housemates, officemates, co-adventurers, co-conspirators, and corner Christmas trees. Thank you for all the memories, balloons, cups of tea, and donuts. Dear lasses, may the road rise up to meet you, the wind be always at your backs, and Santa Claus come to your town in his turbo-sleigh. I am particularly grateful to Dr. Megan Smith for her depth, discipline, and pithy quotes to live by as well as to (Dr.) Laura Ballard for her tall office chair, unimaginable generosity of time and brilliance, carrots, and BNTs notwithstanding the EMSs. FTM.

And last, but certainly not least, I would like to thank my four-legged fellow, Mr. Zeke, for his patience as I pushed the buttons on the magic box, his presence to remind me that we all need treats and walks, and his never-ending tender loving care.

PREFACE

During the year preceding my entry into the Second Language Studies program at Michigan State University, I lived on an island called Ambon in the far East of Indonesia (Appleby, 2010; Deane, 1979). This island is well-recognized for its spices, muck diving, and nearly decade-long civil and religious unrest (Böhm, 2006). My year in Ambon was the last of three that I spent in Indonesia as an English teacher through U.S. Department of State exchange programs. By my third year in Indonesia, I was a highly proficient second language user of Indonesian and experienced no foreign language anxiety. Despite my excellent communicative competence and thrill to be on a tropical island that boasted treasured spices, untouched coral reefs, and tangy street-side fruit salads, it was – for many reasons – a most difficult year. I will spare my readers of the adventures in the Wild Wild East of Indonesia except for one that provides the framing of this dissertation: infrastructure, or rather, a lack thereof.

My main mode of transportation in Ambon was a borrowed motorbike that I used daily to travel to the university's English department. There were a number of ways to reach my end destination. Thankfully, because I had a motorbike, I was not forced to rely upon the not-so-dependable network of island van taxis or ride the splashy powerboats that jetted across the bay. Instead, my choice was between an hour-long motorbike ride around the edge of the bay, or the car ferry. In theory, the car ferry was the faster route, yet the large volume of commuters utilizing the ferry consequently led to longer wait times. Despite my previous two years of Indonesian rubber-time priming[1] and patience training, these long waits to just cross the bay began to grate on me.

[1] *Rubber time* is a term used in Indonesia to describe the lack of time orientation. That is, starting and ending times of events are very rubbery, or flexible. I was *primed*, or exposed, to this rubber

Clearly, I was not the only victim of this glaring infrastructural challenge. As I sat in line and waited for the next ferry, I would gaze at cement pillars rising up out of the water on both sides of the bay. These pillars were the evidence of plans in place to build a bridge that would stretch across the bay. Then, in my mind, I imagined connecting cables and a deck to those pillars. The bridge would someday revolutionize this island; people and cargo would move from one side of the island to the other with unprecedented efficiency.

The bridge-building initiative had been underway for two years before I arrived in 2013 and was projected to be complete in 2016. During my year of living in Ambon, however, it appeared as if no progress was being made on its construction. The pillars stood as the only evidence of the bridge. In short, I was unconvinced that the bridge would ever be completed and wondered how anyone could be so optimistic about the completion of such a large undertaking.

I left the island of Ambon in Summer 2014, and the image of the unfinished bridge remained with me as a representation of some of the seemingly insurmountable challenges I had faced. Similar to the engineers of the bridge project, this optimistic American had braved a large undertaking herself. Time passed in the MSU Second Language Studies (SLS) Program. I read some books and wrote some papers, and I nearly forgot about the unfinished bridge in Ambon until last Fall (2016), I met one of my previous students from Ambon island who is a Fulbright Foreign language teaching assistant (FLTA) at a local university. After updating me on friends and events on Ambon, he announced that the bridge project was complete, and its impact was revolutionary for the islanders.

I was surprised to hear that the bridge project was complete. Finished. Traversable. Operable. This achievement on the other side of the world offered me hope that, even in a

---

time through my two previous years in Indonesia; however, that did not adequately prepare me for the very rubbery time in the Far East of Indonesia.

country where promises and plans were often delayed until “tomorrow” (maybe this week or month) or “tomorrow-tomorrow” (never), a large undertaking could be completed. It was about time.

On an altogether different island, and in an entirely different domain, it was also about time that another infrastructural issue is confronted: the link between SLA and psychology, and more specifically, psychophysiology. This dissertation is, in essence, a bridge-building enterprise that aims to contribute to the incremental work of establishing solid connections between the fields of SLA and psychology. To be clear, these two academic domains are not now mutually unintelligible or completely inaccessible to one another. Researchers have already begun drawing connections between SLA and psychology, and I am eager to add to this emergent enterprise.

In reality, I imagine a bridge-building project is a complicated matter that requires an interdisciplinary team of architects, engineers, and builders to design a secure and crossable passage to the other side. The foundations must be sunk deep into the ground, the cable stretched appropriately between the two towers, the hangers measured appropriately to reach the deck upon which cars and pedestrians are safe to travel.

In many ways, building a bridge is the work of academics everywhere, and I will employ the bridge metaphor throughout this dissertation. In the literature review (chapters 1 and 2), I will discuss the towers and foundations of SLA and psychology. Next, I will introduce the cable – the argument and hypotheses of the present study – that I will stretch between SLA and psychology. The hangers of the bridge metaphor will be the operationalizations of the constructs, measurements, and analyses that I will introduce in chapter 3. Finally, the deck represents the real observations of the present study; in particular, I want to determine how well the hangers (i.e., operationalizations and methodological decisions) stretch between the cable and the deck; that is, how well the operationalizations aligned with the observed data.

Similar to the thrills and challenges of the Wild Wild East of Indonesia, this dissertation has been an intellectual expedition into the Wild Wild West of the SLA field where few researchers have ventured (Dewaele, 2005b; Dewaele, Petrides, & Furnham, 2008; Dörnyei, 2009; MacIntyre, 2002; MacIntyre & Gardner, 1989, 1994b; Schumann, Crowell, Jones, Lee, & Schuchert, 2014). My SLA background and coursework has helped with the foundation of one side of the bridge, and I will describe foundational assumptions and principles necessary to understand the other, the realm of psychology. Since its conception in a psycholinguistics class project, this dissertation truly is exploratory bridgework. My colleague and I set out to investigate the impact of emotional mood induction on adult second language (L2) vocabulary learning. Because the course assignment was *only* a research project proposal, my fellow first-year colleague and I designed an ambitious study that employed a variety of novel measures to the field of SLA: working memory capacity scores, emotional self-reports, three different psychophysiological measures, and four anxiety questionnaires, among other metrics (Fox, 2016; Fox & Miller, 2015; Fox, Miller, Godfroid, & Moser, in preparation; Miller, 2016). An opportunity for summer research funding came along, so my colleague and I submitted this proposal with the understanding that these awards usually went to upperclassmen. To our pleasure and dismay, we received the summer research funding. Our rookie optimism encountered a large undertaking, and we set out to make our grand proposal into a realistic research project.

One of the first steps of this journey was locating the tools and human resources necessary to proceed. One tool we needed was a biofeedback machine that captured galvanic skin responses (GSRs) and heart rate (HR). Biofeedback? Galvanic skin responses? What did those words even mean? It became quickly apparent that my partner and I would need to acquire an entirely new language of terms, research protocols, and computer programs. In addition to locating a GSR machine and acquiring a new language, we sought a more competent peer or expert on the matters of emotions and cognition. One spring afternoon, we left our familiar Wells Hall behind and walked over to the psychology building in search of a collaborator, or rather, a faculty member involved in emotion, cognition, psychophysiology, and/or mood induction research. We identified one faculty member from the hallway bulletin board who matched our criteria, located his office, knocked on the door, interrupted a meeting with his graduate students, and asked for an appointment with him. And the collaboration began.

That afternoon marked the beginning of a collaboration that has lasted nearly two years during which Dr. Jason Moser and his graduate students have generously introduced us to the language of emotions, psychophysiology, and computer programming. Without the cooperation, kindness, and hospitality of the clinical psychophysiology lab and its inhabitants, the "other side" of the bridge would be missing, and this dissertation would not exist. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The purpose of this dissertation is to construct a conceptual argument that bridges emotion with cognition by employing theories and methodologies of second language acquisition and psychology. I am particularly interested in investigating the following variables from the fields of second language acquisition (SLA) and psychology: individual differences, the neurobiology of learning, mood induction paradigm, emotion, and psychophysiology (Clark & Fiske, 1982; Izard, 1989; Leichtman, Ceci, & Ornstein, 1992; Mathieu, Tannenbaum, & Salas, 1992; Schumann et al., 2014; Warr & Downing, 2000).

This dissertation is the culmination of a collaborative research project that began as a psycholinguistics class project. In brief, two groups of thirty-five adult language learners were placed in a negative experimental group or a neutral comparison group and exposed to a series of mood-inducing video-only film clips (Carvalho, Galdo-Álvarez, & Gonçalves, 2012) after which they learned the forms and meanings of 24 Indonesian concrete nouns. Participant physiological response measures (heart rate, heart rate variability, and skin conductance levels) were collected during baseline and film-viewing periods; additional data collected included periodic emotional self-reports, performance on immediate and delayed lexical acquisition posttests, working memory capacity tests (operation span, symmetry span, and non-word repetition [Baddeley, 1992, 2003; Baddeley et al., 1998; Foster et al., 2015]) and a battery of anxiety questionnaires.

As mentioned before, these data were collected in collaboration with another student in my doctoral program. Together, we designed the study, recruited participants, collected the data, and then divided the data analysis into two domains: cognitive (i.e., vocabulary posttests and working memory capacity) and emotional (i.e., anxiety questionnaires and psychophysiology. I selected the analysis of the emotional data; I wrote a qualifying research paper on the results of factor analyses of the four anxiety questionnaires, and this dissertation completes the description of the emotional landscape using physiological data. My colleague selected the cognitive domain and completed his qualifying research paper on those topics; for that reason, I do not attend to the analysis of working memory capacity tests in the present study (Fox & Miller, 2015; Fox et al., in preparation; Mathieu et al., 1992; Miller, 2016; Miller, Fox, Moser, & Godfroid, 2017). General findings of my colleague's study confirmed predictions that exposure to the negative films had a detrimental impact on participants' subsequent vocabulary learning but did not change working memory scores from pre to post mood induction (Miller, 2016; Miller et al., 2017).

Results from the investigation of the role of individual differences in self-reported trait anxiety levels (as determined by four questionnaires) on vocabulary learning (Fox, 2016; Fox et al., in preparation) resulted in no significant influence from the well-known construct of foreign language anxiety; instead, the somatic distress anxiety factor (i.e., physical symptoms such as sweaty palms, racing heart, nausea [Wolfe, 2016]) predicted performance on two subsequent posttests. Participants who reported higher levels of somatic distress (1 SD above the mean) and were exposed to the neutral film clips performed significantly better on the productive and receptive vocabulary posttests than those who reported average or below-average levels of somatic distress. The effect of this factor score was wiped out, however, for those in the negative mood induction group; the impact of the negative mood induction resulted in poorer performance on the posttests overall, regardless of somatic distress levels. Of the seven possible anxiety factors, including foreign language anxiety, it is noteworthy that the somatic distress factor uniquely predicted performance on two of the vocabulary posttests.

These findings appropriately set the scene for the present study and its investigation of somatic distress and physiological responses to mood induction. In this dissertation, I analyzed the physiological data from the *hot cognition* (Brand, 1987) study to explore how language learners' physiological responses to emotional stimuli impact their learning in a paired-associates vocabulary-learning task.

As introduced in the preface of the dissertation, I am employing the architecture of a bridge throughout the dissertation in order to provide a framing for the interdisciplinary project. It is my aim to build off of SLA and psychology theory and operationalizations in the present study. The purpose of this first chapter is to establish the tower of second language acquisition (SLA) on one side of the river. I will first describe the onramp of this tower, foreign language anxiety, how it relates to the present study, and why it is particularly important for the field of SLA. Next, I will situate foreign language anxiety within a larger investigation of emotions and cognition in SLA. Then, I will provide a rationale for employing a paired-associates vocabulary-learning task as the operationalization of cognition within SLA.

In the second chapter, I will present the second tower of interest, that of psychology and psychophysiology. The onramp of this chapter is the construct of anxiety within psychology and how it is defined, elicited, and measured. Next, I will introduce *psychophysiology*, including the assumptions and caveats necessary to understand how researchers link mechanisms of the mind and body. I then posit that emotions are connected to an individual's cognition, attention, and learning. By extension, I argue that emotions can influence the outcomes of a paired-associated learning task, which will be used to represent cognition. I conclude chapter 2 with an overview of the constructed towers of chapters 1 and 2, review the most pressing of research questions, and offer conceptual research questions and hypotheses for the present study.

In summary, the overarching purpose of the present dissertation is to investigate to what extent emotion–as measured by psychophysiology and self-report–during an emotional induction experiment influences subsequent paired-associates vocabulary in a novel language. While there is considerable interdisciplinary research at the intersection of emotion and cognition within SLA (Arnold, 2011; Dewaele, 2005b; Dörnyei, 2009; Douglas Fir Group, 2016; Izard, 1989; Swain, 2011), the present study is arguably the first to employ measures of physiology to SLA outcome variables. This is an exciting new frontier within SLA, and as I set out on this bridge-building enterprise, I am particularly motivated by the words of Dewaele (2005b):

> There will always be enough space for academic border-crossers to plant new flags in "unclaimed territories"…I fervently believe that a stronger focus on physiological, psychological, affective, and emotional issues in SLA can provide crucial theoretical insights into L2 acquisition that are now missing (p. 577-8).

CHAPTER 1

Tower 1: Second Language Acquisition

Emotion and SLA

Research on the intersection of emotions and SLA is not an altogether novel concept; indeed, there have been many studies investigating different aspects of the link between emotion and cognition. The majority of research on emotions in SLA has pertained to emotionally-laden vocabulary: the mental representation of emotional vocabulary in bilinguals (Altarriba, Bauer, & Benvenuto, 1999; Altarriba & Canary, 2004; Pavlenko, 2008); L2 taboo-word sensitivity (Eilola, Havelka, & Sharma, 2007; Harris, 2004; Harris, Aycicegi, & Gleason, 2003); swearing in participants' L1 versus L*n* (Dewaele, 2004a, 2004b, 2005b); and pedagogical considerations of teaching emotionally-laden L2 materials (Dewaele, 2005a; Dewaele & Pavlenko, 2002; Gregersen & MacIntyre, 2014; Gregersen, Macintyre, & Meza, 2014; MacIntyre, 2012; Oxford, 2014; Young, 1994). There is growing attention to the role of student and instructor affect in the language (Arnold, 2011; Krashen, 1985; Krashen, Dulay, & Burt, 1982; Schumann, 1999), with the bulk of studies dedicated to foreign language anxiety (Horwitz, Horwitz, & Cope, 1986; MacIntyre & Gardner, 1989). Due to its predominance in research and greatest potential for pedagogical implications, I employ foreign language anxiety as the onramp to the present study.

Foreign language anxiety. Foreign language anxiety is defined as the worry and negative emotional reaction aroused when learning or using a second language (MacIntyre & Gardner, 1994b). Previous researchers have agreed that foreign language anxiety is a situation-specific anxiety and arises during tasks, such as speaking (MacIntyre, 2012), reading (Matsuda & Gobel, 2004) or during particular phases of the learning process (MacIntyre & Gardner, 1994b). However, what remains inconclusive is the impact of foreign language anxiety on language learners. Previous reports on foreign language anxiety have uncovered conflicting evidence about whether foreign language anxiety is a facilitating (MacIntyre, 1995) or debilitating (Ehrman & Oxford, 1995; MacIntyre, 2002; Robinson, 2002) construct in the language-learning context. While foreign language anxiety has been reported to be consistently more present among less successful learners (Horwitz, 2001; Horwitz & Young, 1991; Horwitz, 2010), "facilitating anxiety" (Spielmann & Radnofsky, 2001, p. 262) can also serve as a "helpful wake-up call" (Oxford, 2014, p. 595) to energize learners and usher them into a productive state of alertness (Alpert & Haber, 1960; Kleinmann, 1977; Spielmann & Radnofsky, 2001, p. 262). Such inconclusive results concerning foreign language anxiety may lead one to question the underlying, latent variable(s) involved. I now turn to questions surrounding the construct of foreign language anxiety: its measurement, underlying constructs, and influence within the field of SLA.

Measuring foreign language anxiety. Considerable work has been done to quantify and qualify the existence of foreign language anxiety, and researchers have employed various methodologies ranging from qualitative case studies (Bailey, 1996; Benesch, 2013; Bown, 2009; Oxford, 2014) to quantitative path analyses (MacIntyre & Charos, 1996). The vast majority of reports on foreign language anxiety have utilized mixed methods involving stimulated recall, interviews, and self-report questionnaires, such as the State-Trait Personality Inventory (Spielberger, 2010), and the Anxometer (MacIntyre & Gardner, 1994a). Most recently, the idiodynamic method, a stimulated recall of anxiety levels during an oral presentation, has been introduced as another way to quantify foreign language anxiety (Gregersen et al., 2014; MacIntyre, 2012; Macintyre & Legatto, 2010). By far, the most widely employed measure is the Foreign Language Classroom Anxiety Scale (Horwitz et al., 1986), a questionnaire of 33 statements regarding participants’ study of a foreign language presented with a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). This scale defines foreign language classroom anxiety as “a distinct complex of self-perceptions, beliefs, feelings and behaviors related to classroom learning arising from the uniqueness of the language learning process" (Horwitz et al., 1986, p. 128).

Despite its apparent authority on the construct of foreign language anxiety in SLA, the FLCAS has received criticism regarding its internal validity and state versus trait dimensionality (Young, 1994). Horwitz et al. (1986) advocated that the items on the FLCAS reflect communication apprehension, test anxiety, and fear of negative evaluation in the foreign language classroom, other researchers have identified various underlying constructs in the questionnaire using factor analyses.[2] MacIntyre and Gardner (1989) uncovered two factors underlying the FLCAS, which accounted for 48% of the variance: general anxiety and language anxiety. Later, MacIntyre and Gardner (1991) identified three dimensions: fear of negative evaluation, communication apprehension, and test anxiety (Arnaiz & Guillén, 2012; Liu & Jackson, 2008). These findings call into question the uni-dimensionality of the FLCAS due to the fact that researchers have consistently identified more than one construct within FLCAS. Another concern about the questionnaires is whether the FLCAS represents a more permanent, trait-level or a more transient, state-level, individual difference. Despite the fact that researchers have confirmed foreign language anxiety has no relationship with trait-level anxiety (Gregersen et al., 2014; Horwitz et al., 1986; MacIntyre & Gardner, 1989), this measure continues to be used as an operationalization for language learners’ trait-level self-evaluations of foreign language

---

[2] Factor analysis is a statistical method of determining the number of subscales, or factors, a questionnaire may have based upon how each individual questionnaire item loads onto a factor. This method is particularly helpful in developing questionnaires and reducing the amount of data (Pett, Lackey, & Sullivan, 2003).

[3] A brief search on Google Scholar using the keyword search “foreign language classroom

anxiety. Taken together, previous analyses that have employed the FLCAS as a uni-dimensional, trait-level measure may require careful reconsideration.

It is somewhat disconcerting that such a large number of researchers have employed this scale based upon the inconclusive assumption that foreign language anxiety is a singular trait, and further, operationalized foreign language anxiety as a trait-level characteristic. To date, countless studies[3] have investigated the relationship of foreign language anxiety measured by the FLCAS instrument (Horwitz et al., 1986) on SLA performance variables such as the following: foreign language reading anxiety (Matsuda & Gobel, 2004); course grades (MacIntyre & Gardner, 1994a); self-efficacy (Mills, Pajares, & Herron, 2006); standardized language proficiency scores (Dewaele, 2013; Marcos-Llinás & Garau, 2009); cognitive, affective, personality, and demographic variables (Onwuegbuzie, Bailey, & Daley, 2000); number of additional languages (Dewaele, 2002; Dewaele, 2007); willingness to communicate, communicative competence, and communication apprehension (MacIntyre, 1994); and general learning anxiety (Sparks & Ganschow, 2007).

Expanding anxiety in SLA. Due to the abovementioned concerns of the inconclusive construct foreign language anxiety being correlated with many variables in SLA and psychology, in an earlier study I investigated the underlying constructs of the FLCAS and three additional psychological anxiety questionnaires. My aim was to uncover a more comprehensive and nuanced understanding of the construct of anxiety for foreign language anxiety and the other anxiety questionnaires, and in turn, be able to make more realistic and accurate interpretations of results (Fox, 2016; Fox et al., in preparation). The four anxiety questionnaires underwent separate exploratory factor analyses and yielded seven underlying factors. The FLCAS

[3] A brief search on Google Scholar using the keyword search "foreign language classroom anxiety scale" produced a modest 84,400 results.

questionnaire produced two underlying factors: foreign language anxiety and general learning anxiety. In short, this study revealed that anxiety, and foreign language anxiety specifically, is not as straightforward as one may hope. Except for one vocabulary posttest, the two factors of the FLCAS did not provide explanatory power in the vocabulary posttests. Rather, the somatic distress (i.e., physical symptoms such as sweaty palms, racing heart, nausea [Wolfe, 2006]) factor score impacted vocabulary performance across two vocabulary posttests. Those who reported higher levels (1 SD above the mean) of somatic distress and were exposed to the neutral film clips, performed significantly better on the productive and receptive vocabulary posttests than those who reported average or below-average levels of somatic distress. However, there were no significant effects of this factor score in the negative group, and we hypothesized that the negative mood induction wiped out any effect of this factor. The results of this study align with previous literature marking the compensatory benefit of increased anxiety levels on more complex tasks. Because of their awareness of their reduced processing effectiveness, participants pay increasingly more attention to the stimulus (Eysenck, 1979; MacIntyre, 1995).

While some may find it unreasonable to imagine that differing levels of self-reported somatic distress have an impact on a subsequent vocabulary-learning task, there is mounting evidence that such a connection exists. In fact, it is this link which I investigate in the present study: the conceptual bridge linking the constructs of emotion and cognition. In particular, I am interested in investigating the subconscious, autonomic representations of emotion on vocabulary learning through the paired-associates learning task. To date, with the exception of Gregersen et al. (2014), studies within the field of SLA have not explored psychophysiological reactions. The purpose of the present study is the fill the gap of understanding of how anxiety—and emotion more generally—can be best measured and observed. The specific details concerning sympathetic nervous system responses will be further explained in chapter 2, and I will now review the few SLA studies that have incorporated aspects of emotional mood induction or physiological measures.

Mood induction and SLA. One of the earliest attempts to measure environmentally manipulated emotion involved a researcher who interviewed participants in either an anxiety or a non-anxiety-inducing condition (Steinberg & Horwitz, 1986). Those in the anxiety condition were notified of the presence of a video recorder that was "conspicuously" (p. 131) played by the researcher, who treated the participant "brusquely" and maintained a "cold and official posture" while providing "stress-loaded" instructions that emphasized that "good performance was crucial to the success of the experiment." Participants were then asked to describe a picture. Results indicated that those in the negative experimental condition described visual stimuli less fully than the comparison group; based on their results, the researchers claimed that constant feelings of pressure and fear of negative evaluation in the classroom had detrimental impacts on language performance.

Second, MacIntyre and Gardner (1994b) aimed to test the theory of *negative interference* (Tobias, 1986). Tobias defined the term as negative thoughts related to one's failure, self-deprecation and avoidance, and further argued that *negative interference* had the potential to interfere with three phases of learning – input, processing, and output phases – by monopolizing cognitive resources from separate "fairly intense cognitive activities" (Khan & Zafar, 2010; MacIntyre & Gardner, 1994a, p. 2), such as language learning. Tobias' concept of negative interference has been well-cited in SLA research concerning foreign language anxiety and associated with similar constructs such as high affective filter (Krashen, 1985), low willingness to communicate (MacIntyre, 2007), and a high foreign language anxiety (Horwitz et al., 1986).

In the study by MacIntyre and Gardner (1994b), the selected anxiety-eliciting device was a video camera introduced at differing stages of the experiment which involved a paired-associates lexical-learning task. Results indicated that the presentation of the video camera at any phase of the experiment increased participant anxiety as well as decreased working memory and language task performance.

Physiology and SLA. Gregersen et al. (2014) pioneered the frontier of physiology and SLA. They collected continuous heart rate data of six language learners during a foreign language class oral presentation, in combination with their self-developed Anxometer, an idiodynamic measure of state anxiety, and an adapted version of the Foreign Language Classroom Anxiety Scale (Gregersen et al., 2014). Their research questions focused on the change in heart rate over time and whether these heart rate changes correlated with phase and trait-level self-reports of anxiety, represented by the Anxometer and FLCAS, respectively. Although this study introduced novel methodological applications to explore the relationship among emotions, physiology, and language, two areas of concern rise to the surface: the way that heart rate was operationalized as the index of not only arousal, but also it's being linked with the more specific emotion of foreign language anxiety. As will be further explained in chapter 2, the investigators established some far-reaching connections between psychophysiology and emotions resulting from the oral presentation.

Other physiologically-focused research concerning mood induction and SLA has focused on the processing of syntactic anomalies (Croft, Gonsalvez, Gander, Lechem, & Barry, 2004; Witt et al., 2006) and semantic infelicities (Verhees, Chwilla, Tromp, & Vissers, 2015) observed through ERP error patterns. Similarly to the present study, these studies employed a mood induction procedure (MIP) through film clips, which were intended to elicit sad and happy moods. Both studies revealed that participants exposed to the happy film clips exhibited qualitative superiorities in performance of subsequent linguistic tasks. Chwilla, Virgillito, and Vissers (2011) theorized that emotions do impact processing styles: a positive mood leads to greater cognitive flexibility, and a negative mood may "focus our attention more narrowly on specific details of a situation [which] invalidates accessible cognitions and fosters local, item-specific processing" (p. 2411). Although this novel line of research utilizes advanced methodologies to investigate the role of emotions on linguistic processing, there remain a number of caveats. First, while the researchers reported successful manipulation checks of the film clips, the film clips were not normed on a larger population, standardized in length, or designed to elicit a specific emotion. The self-report measure targeted participants' emotion responses to participants' perceived happiness and sadness levels, which some may argue, limits the scope of potentially elicited emotions. Literature from the field of psychology cautions researchers that targeting specific emotions in experimental tasks and later making more general interpretations based upon those discrete emotions may lead to inaccurate conclusions (Berntson & Stowell, 1998; Gross & Levenson, 1995; Lang, Greenwald, Bradley, & Hamm, 1993; Mauss & Robinson, 2009; Plutchik, 1980). What is preferred instead is a more dimensional approach to emotion, which will be reviewed in chapter 2.

In summary, extant literature provides the foundation of our understanding regarding the relationship between emotion and cognition in the field of SLA. Still, there remains much to be discovered regarding the impact of mood induction, represented by physiology measures and self-report, on language-learning outcomes. The next section introduces the cognitive perspective of the present study, namely how anxiety can be understood as a type of cognitive interference to attentional processes required for successful language-learning. I will introduce the linguistic task and feature that will be employed in the present study, namely intentional vocabulary learning represented by a paired-associates vocabulary-learning (PAL) task. Then I will conclude this chapter with a note about how the PAL may inform early processes of SLA.

Emotion and cognition

Much of our understanding of how anxiety affects SLA—and more generally, how emotion is linked to cognition within SLA—is owed to the expansive and interdisciplinary work of MacIntyre, Gardner, and Schumann (MacIntyre, 1994, 1995, 2002; MacIntyre, 2012; MacIntyre & Gardner, 1989, 1994a, 1994b; Schumann, 1997, 1999; Schumann et al., 2014). These early pioneers of the SLA-psychology frontier were not only attuned to, or aware of, foreign language anxiety; in addition, they actively sought converging evidence in the neighboring fields of psychology, evolutionary biology, and cognitive science to understand how foreign language anxiety relates to basic human responses to threat. For instance, in his 1995 article, MacIntyre examined the impact of anxiety upon two dimensions–the behavioral and the cognitive. In the domain of behavioral consequences of anxiety, he noted that anxious participants exhibited increases in sympathetic nervous system arousal and attempt to escape or leave a threatening situation (Levitt, 1980; MacIntyre, 1995; Sarason, 1986; Spielberger & Spielberger, 1966). This relates to the psychological notion of *stimulus appraisal*, which was introduced to the field of SLA as a partial explanation of the roles of affect and motivation in language acquisition (Schumann, 1997, 1999; Schumann et al., 2014). Schumann referenced Scherer and colleagues' (1984) five categories of stimulus appraisal and related those to how language learners assess the environment of a language-learning classroom:

> …whether the situation is *novel* or *pleasant*, whether it contributes to one's *goals or needs*, whether we feel we have the *coping potential* to deal with its consequences, and how our engagement with a situation may affect our *self and social image.* (Mates & Joaquin, 2013, p. 421)

Such a concept pairs well with the construct of foreign language anxiety in that those individuals who feel an element of threat, internally (negative self-perceptions) or externally (fear of negative evaluation), may respond with anxious behaviors that lead to less than ideal levels of engagement, social interaction, and risk-taking often requisite in language-learning contexts. Furthermore, it highlights the ultimate variability of individual differences among language learners. Due to each language learner's unique stimulus appraisal system, we can conclude that individuals will also respond to various language learning situations uniquely (Mates & Joaquin, 2013).

Just as MacIntyre and Dewaele drew from other fields to explain emotional foundations of foreign language anxiety, they also looked to other disciplines for links between SLA and cognition. MacIntyre (1995) reasoned that the role of anxiety in second language learning, based on the study of the psychological effects of anxiety, could be understood in this way: "language learning is a cognitive activity that relies on encoding, storage, and retrieval processes, and anxiety can interfere with these by creating a divided attention scenario for anxious students" (p. 91). The assertion that anxiety interferes with the learning process aligns well with three-part model for the role of attention in SLA proposed by Tomlin and Villa (1994). Attention was defined as a cognitive process that requires individuals to sort through overwhelming amounts of sensory and cognitive information called L2 input. The three components of attention introduced by Tomlin and Villa were: *alertness, orientation,* and *detection*. In their model, *alertness*, comparable with stimulus appraisal (Schumann, 1999), is defined as "an overall readiness to deal with incoming stimuli" (Leow, 1998, p. 135); *orientation* is the commitment of attention resources to stimuli, which may facilitate or have no impact on the final phase of *detection*, the "cognitive registration of the stimuli" (Leow, 1998, p. 135). In this three-component model of attention, the authors noted that success in SLA ultimately requires detection, and although alertness and orientation are not required for detection to occur, those two factors can serve as facilitation for detection. According to the authors, attention was considered to be a limited capacity system that requires individuals to select information for further effortful, or intentional, processing; such controlled cognitive processing, they noted, requires attention, and thus creates competition with other processes requiring attention.

Orienting will play an important role in the present study, so it is important to note that this "specific aligning of attention on a stimulus has facilitative or inhibitory consequences for further processing depending on whether information occurs as expected or not as expected" (Tomlin & Villa, 1994, p. 191). Although the orientation effect that I observed in the present study occurred during the mood induction, and not during the vocabulary-training phase, I suspect that this orientation, in conjunction with the mood induction and other individual differences, had an impact on vocabulary learning.

Paired-associates vocabulary learning (PAL). In the present study, I employ a paired-associates vocabulary-learning task as the operationalization of cognition and attention. Compared to its counterpart, incidental learning, researchers claim that intentional, or deliberate, learning is a more efficient and convenient process of learning in that it allows learners to consciously focus on aspects of word knowledge, and control the repetition and processing of the vocabulary to secure learning (Hulstijn & Laufer, 2001; Laufer, 2005).

The PAL paradigm is the laboratory equivalent to a language classroom wordlist exercise: a new foreign word is presented, or paired together, with either the learners' native language (L1) translation or a picture in isolation. PAL has been employed in previous research because it allows for controlled exposure to specific aspects of language (Folse, 2004), and other researchers confirmed that L1/L2 word pairs are particularly helpful in the early stages of vocabulary learning because they capitalize on learners' L1 (Lotto & de Groot, 1998; MacIntyre & Gardner, 1994b). In his book on myths of vocabulary learning, Folse (2004) rejected the myth that the use of translations to learn new vocabulary words should be discouraged and promoted the use of word lists in foreign language classrooms as an opportune method for intentional learning of a limited number of words. Similarly, previous researchers have noted that learning the L2 word/picture pair is particularly productive because of the dual encoding of the verbal and imagery information in the brain. When pitted against one another, though, researchers found that the L1/L2 word pair results in better retention than L2 word/picture pairings (Paivio and Desrochers (1981), Lotto and de Groot (1998). As will be noted in the methodology section, the present study paired both L1 translations and pictures along with the new L2 word.

My rationale for operationalizing cognition through a vocabulary-learning task is threefold. First, a vocabulary-learning task was chosen because it provided materials that were discrete, controlled, lab-based equivalent to word lists in the language learning classroom. Next, within the diverse landscape of first and second language acquisition theories, vocabulary learning is generally understood to represent a domain-general type of learning and cognition (Fedorenko, 2014). The final reason that the paired-associates vocabulary-learning task was selected as the outcome variable in the present study is that vocabulary learning is one of the most critical initial steps, and ongoing quest, in studying a language. Vocabulary acquisition is an intricate and "complex phenomenon" (Schmitt, 2014, p. 914) because each word, or word family, is composed of multiple depths or qualities and breadths or sizes. Vocabulary depth refers to the multifaceted knowledge about a word, of which some argue there are up to seven types: phonological, orthographic, syntactic, morphological, pragmatic, articulatory, idiomatic, and semantic (Alderson, 1984; Anderson & Freebody, 1981; Beck, Perfetti, & McKeown, 1982; Coady, Magoto, Hubbard, Graney, & Mokhtari, 1993; Davis, 1944; Folse, 2004; Kameenui, Carnine, & Freschi, 1982; Koda, 1989; Laufer, 1992; Laufer, 1997; Nation, 1993; Nation & Coady, 1988). Vocabulary breadth is the number of words an individual knows receptively and productively (Elgort & Nation, 2010; Folse, 2004; Laufer, 2005; Schmitt, 2008). Despite the apparent complexity of vocabulary acquisition of a novel language, Lewis (1993) advocated that vocabulary knowledge is the "core or heart of language" (p. 89). Similarly, Wilkins (1972) famously stated, "while without grammar very little can be conveyed, without vocabulary nothing can be conveyed" (p. 111). Correlational studies have further boosted its prestige by revealing strong relationships between vocabulary knowledge and global assessments of writing quality (Folse, 2004; Green & Meara, 1995; Meara, 1980) as well as reading comprehension in individuals' L1 and L2 (Astika, 1993; Engber, 1993). In sum, vocabulary is a critical component of language acquisition that makes it a rich and tangible research domain for investigation.

Taking a step back, it is understandable that successful L2 vocabulary acquisition, in all its necessary depth and breadth, will not occur by learning words from lists. Researchers and language instructors alike agree that vocabulary learning from lists is inadequate and superficial (Schmidt, 1992). Accordingly, PAL, considered to be the laboratory-based equivalent to word lists, cannot be generalized to larger constructs of L2 vocabulary acquisition. Thus, it would be a gross overestimation to posit that participants who participate in a PAL task and learn 24 concrete nouns have acquired a new language, let alone vocabulary. Nevertheless, there is worth in the present study and its operationalization of cognition through the PAL. Despite the diversity and complexity of second language vocabulary learning, the PAL will provide a glimpse into the effect of intentional learning of discrete-items on a word list, and this, in turn, may have implications for instructed SLA.

Instructed SLA. Although it may seem premature to suggest implications of the present study, there is reason to believe that language instructors may benefit from its results. Although some would argue that second language acquisition is primarily the responsibility of the language learning individual, the language instructor plays a critical role in the explicit instruction and type of input she provides to the learner. Undoubtedly, if learners are to be explicitly taught all necessary vocabulary items for survival alone, the instructors may feel particularly burdened by such a mountainous task (Ellis & Beaton, 1993). Debate continues surrounding the exact proportions of explicit versus implicit vocabulary instruction; nevertheless, literature on instructed L2 vocabulary acquisition proposed that the role of conscious noticing and explicit focus on vocabulary form and meaning are effective (Doughty, 1991; Lightbown & Spada, 1990; Long, 1983; Schmitt, 2008; Tomlin & Villa, 1994), and particularly for new learners, new words should be presented out of context through tools such as word lists (i.e., a vocabulary flood, [Clipperton, 1994; Laufer, 2005]). Schmitt recommended that instructors take an incremental approach to vocabulary instruction that corresponds with the various stages of language proficiency: beginner learners ought to be taught through explicit methods to establish the form-meaning link, and higher proficiency learners should be encouraged to develop strategy use, word repetition practices, and a long-term commitment toward holistic mastery (Nation, 2013). How exactly – the frequency and exposure method – vocabulary items ought to be taught in the language-learning classroom are outside the scope of the present study. However, what may prove to be insightful for language instructors is how varying levels of anxiety and environmental factors may impact subsequent explicit, or intentional, learning in the classroom.

Summary

The first chapter has established the SLA perspective on foreign language anxiety and emotions, reviewed previous literature concerning mood induction and physiological measures, and introduced the rationale for the vocabulary-learning task as the operationalization of intentional learning and cognition. There are a few highlights from this chapter that bear repeating. First, foreign language anxiety is only one aspect of emotion in SLA, and there are many opportunities to expand and develop our field's understanding of, elicitation, and awareness of individual differences related to internal and external factors of language learners (Mates & Joaquin, 2013). One framework we can employ to examine the impact of emotions on cognition in SLA Tomlin and Villa's (1994) 3-component attention model involving: alertness, orientation, and detection. As previously described, attention is a critical component of second language acquisition, and the orienting effect proves to be of particular importance in the present study. Finally, I presented the PAL task as the operationalization of cognition, attention, and more specifically, initial stages of intentional/explicit learning in SLA in the present study (MacIntyre, 2002; Robinson, 2002). I now turn to the field of psychology and its approach to the study, elicitation, and measurement of emotion. In particular, I will highlight the assumptions and metrics involved in psychophysiology.

CHAPTER 2

Tower 2: Psychology

Emotion and psychology

Much like the onramp for the SLA tower in chapter 1, I begin this chapter with a review of the construct of anxiety from the vantage point of the field of psychology. Anxiety is defined as the subjective feeling of tension, apprehension, nervousness, and worry associated with an arousal of the autonomic nervous system (Spielberger, 2010). This feeling is expressed uniquely across individuals, and symptoms range from worry, somatic distress, cognitive distress (i.e., feelings of impending doom, low self-efficacy, and low engagement), to even depression (Seidenberg, Haltiner, Taylor, Hermann, & Wyler, 1994). In addition to the various faces of anxiety, duration of symptoms also vary: trait, state, and situation-specific (MacIntyre & Gardner, 1989). Trait anxiety is consistent over time and context (i.e., commonly associated with the personality trait known as neuroticism [Dewaele, 2002, 2013]). State anxiety is a more temporary experience of anxious symptoms, usually due to anxiety-provoking experiences. Situation-specific anxiety pertains to specific situations that arouse anxiety, such as public speaking (Smith, Sawyer, & Behnke, 2005; Witt et al., 2006), math, test-taking (Fresco, Heimberg, Mennin, & Turk, 2002; Zeidner, 1998), and not surprisingly, the foreign language classroom (Horwitz et al., 1986; MacIntyre & Gardner, 1994b; Schmader & Johns, 2003). In sum, anxiety is a multifaceted construct.

Yet despite the apparent intricacies, anxiety is situated within the even more elaborate landscape of emotion. In both the academy and the real world, emotions are complicated. Kreibig (2010) defined emotion as a "multi-component response to an emotionally potent antecedent event, causing changes in a subjective feeling quality, expressive behavior, and physiological activation" (p. 397). This definition underscores the complexity of emotion even in a sterile and controlled laboratory setting. An added layer of complexity involves the emotionally potent antecedent event that stimulates or activates an emotional response. This too, no doubt, is inherently variable across individual, time, and space. Individuals do not universally respond to emotional events in the same manner and time frame, or with the same intensity. In sum, the study of emotion is a delicate matter, and this review will focus in particular on the naming, elicitation, and measurement traditions within psychology.

Classifying emotion. Everyday language for emotion includes words such as fear, anger, love, surprise, and sadness; some psychologists believe that there is a universal set of distinct, or primary emotions such as fear, disgust, and anger (Gross & Levenson, 1995; Plutchik, 1980). For instance, a study by Izard (1997) identified ten basic emotions: joy, interest-excitement, surprise, sadness, anger, disgust, contempt, fear, shame, and guilt. More recently, love (Shaver, Morgan, & Wu, 1996) and pride (Tracy & Robins, 2004) were also proposed as discrete emotions. Previous research has linked these discrete emotions with specific physiological measures (Ekman, Levenson, & Friesen, 1983; Levenson, 1992; Levenson, Ekman, & Friesen, 1990; Sinha, Lovallo, & Parsons, 1992; Stemmler, 1989; Witliet & Vrana, 1995).

Despite the apparent plausibility and previous findings of isolating and measuring pure emotions, not all psychologists agree with the discrete approach to emotions (Christie & Friedman, 2004; Mauss & Robinson, 2009). An equally valid approach to classifying emotion is through a dimensional understanding in which three variable emotional dimensions of *valence*, *arousal*, and *dominance* are used to describe a wide range of emotions (Bolls, Lang, & Potter, 2001; Lang et al., 1993). In brief, the psychological construct of *valence* is defined as the range from sadness (low valence) to happiness (high). *Arousal* is the range that an individual is "wide awake, alert, vigorous, and full of pep" (Revelle & Loftus, 1992, p. 115) versus being relaxed, sluggish, and tired. *Dominance* is the extent to which an individual feels dominated and submissive or dominant and in control. The Circumplex model of emotions (Lang, 1979; Lang & Bradley, 2007; Levenson et al., 1990; Mauss & Robinson, 2009), suggests that valence and arousal can be used as the two axes for plotting the general landscape of emotion (Bolls et al., 2001). Valence runs horizontally, and arousal stretches vertically. For example, anger could be described as an emotion with low valence, high arousal, and high dominance, and joy is the combination of high valence and high arousal. Anxiety, on the other hand, does not appear in the model. If it did, it would be situated in the upper right quadrant, revealing a combination of low valence and high arousal. In summary, the study of emotion is intricate and complex, resulting in methodological challenges of controlling for individual differences as well as eliciting true-to-life results in a laboratory setting. Baumeister, Vohs, and Funder (2007) label this challenge as lab-based pseudo behaviors that may not reflect what would happen in the real world, further underscoring the difficulty for theory-building and -testing within the domain of emotion research. The next section of this chapter describes psychological approaches to eliciting emotions within a laboratory setting.

Eliciting emotion. *Hot cognition*, first coined by Brand (1987), was the initial term for experimental paradigms that examined how changes in emotional states impacted cognitive tasks. Researchers that employed this methodology manipulated standard "cold," unemotional experimental conditions to observe how participants responded to mood induction tasks. The term *hot cognition* is outdated and has been replaced with the terms *mood induction* or *manipulation* in the field of psychology. For that reason, I employ the term mood induction in the present study.

Experimental psychology has used many methodologies to elicit emotion: imagery, threat of electrical shock, personal recall, startle reflexes, musical excerpts (Etzel, Johnsen, Dickerson, Tranel, & Adolphs, 2006), preparation for a speech before an audience (Pauls & Stemmler, 2003), facial expressions (Adolphs, 2002), affective pictures (Lang, Bradley, & Cuthbert, 1990), and film clips (Kreibig, Wilhelm, Roth, & Gross, 2007). In the current study I utilize films clips from an established modern affective movie database (referred to as the Emotional Movie Database, or EMDB, [Carvalho et al., 2012]). Film clips for mood induction research has been a commonplace approach for the following reasons: ease of standardization, lack of need for deception, and ecological validity (Baldaro et al., 2001; Carvalho et al., 2012; Christie & Friedman, 2004; Codispoti, Surcinelli, & Baldaro, 2008; Demaree, Robinson, Everhart, & Schmeichel, 2004; Fernandez et al., 2012; Gross, 1998; Gross & Levenson, 1995; Gross & Levenson, 1993; Kreibig et al., 2007; Lazarus, Speisman, Mordkoff, & Davison, 1962; Palomba, Sarlo, Angrilli, Mini, & Stegagno, 2000; Sternbach, 1962; Vianna & Tranel, 2006; Viinikainen, Glerean, Kettunen, Sams, & Nummenmaa, 2012). In the extant reviewed SLA literature, film clips have yet to be employed beyond the domain of pedagogical and assessment purposes (Jaén & Basanta, 2009; Winke, Gass, & Sydorenko, 2010).

Measuring emotion. As noted in the definition of emotion, the three approaches for operationalizing emotional constructs are behavioral observations, self-reports, and physiological measures (Lang, 1979; Lang & Bradley, 2007; Levenson et al., 1990; Mauss & Robinson, 2009). Researchers recommend collecting a combination of these measures to triangulate multidimensional emotional data (Cowley, Fantato, Jennett, Ruskov, & Ravaja, 2014; Cowley et al., 2016; Mauss & Robinson, 2009).

Behavioral reports. Behavioral reports are trained observers' assessments of a participant's vocal characteristics, facial behavior, and other observable aspects of the individual (Bradley & Lang, 2000; Mauss & Robinson, 2009). Proponents of the observational approach deem this practice to be the purest, superior means of data acquisition in that "what is being observed *is* in fact what is being studied, what is being studied is 'there' as plainly as the eyes can see" (p. 4). Although behavioral observations offer systematic procedures for direct data collection, limitations with this approach exist. First, successful employment of this method requires extensive (and expensive) training and well-normed raters, who may possess particularly influential biases in their decision-making processes. Another potential limitation is the lack of rigorously evaluated and validated coding systems (Reis & Judd, 2000). A final risk of behavioral observations is that complex behaviors can be reduced to a few quantifiable variables, leading to oversimplification of the phenomena of interest (Suen & Ary, 1989). According to Baumeister et al. (2007), the era of direct observation is declining due to these challenges and limitations, and also the general shift toward inner processes of individuals. They wrote:

> Once upon a time, perhaps, psychologists observed behavior and reported what they saw, along with their theories about why it happened. The emergence of competing theories, and therefore competing explanations, led psychologists to push each other to show what happened inside the person to produce the behavior. Gradually the focus shifted on these debates about inner processes, somewhere along the way, it became acceptable to publish data on inner processes without any real behavior included at all, which eventually became the norm. (p. 400)

The present study is a consequence of the shift described above: its foci are inner processes of self-reports and physiological changes, in lieu of behavioral observations.

Self-report. Common to both fields of SLA and psychology, self-report measures are appraisals of one's subjectively experienced emotions collected through think-aloud protocols, interviews, and questionnaires. Despite the value of gaining first-hand insight into a participant's emotional experience, researchers have cautioned using self-reports as the sole source of data collection due to biases such as social desirability (Gregersen et al., 2014; Robinson & Clore, 2002; Srivastava, Guglielmo, & Beer, 2010; Stone, Bachrach, Jobe, Kurtzman, & Cain, 1999; Vazire & Carlson, 2011; Watson & Gatchel, 1979) or the Pollyanna Principle, which is the general tendency to favor positive over negative responses (Boucher & Osgood, 1969; Kihlstrom, Eich, Sandbrand, & Tobias, 2000; Matlin & Stang, 1978).

For example, West and Brown (1975) sought to demonstrate the apparent unreliability of individual self-reported behavioral predictions against the individuals' actual observed behaviors. The researchers conducted the same experiment in two different ways: first, individuals completed a self-report questionnaire of how they would respond in hypothetical situations (i.e., an accident victim standing on the street); next, those same participants were placed into the same hypothetical situation and their behavior was observed. Unsurprisingly, individuals' self-reports revealed more generous and kind responses to the victims of the hypothetical events than when they were observed in reality. That is, there was a disconnect in the way that individuals self-reported their expected behavior compared to how they behaved in the actual staged event. In effect, this study emphasized that caution is necessary in building theories based solely upon self-reports and reinforced the importance of collecting data from multiple sources. Robinson and Sedikides (2009) discovered that self-reports were most reliable when used to report "online," or currently experienced emotions as opposed to questionnaires that tapped into past, future, or trait-related entities (e.g., trait-level anxiety), which may be at risk for errors in perception and memory (Kihlstrom et al., 2000). In the present study, I use self-reports to measure both trait and state level emotional changes.

Psychophysiology. The physiological measurement of emotion is the focus of the present study. Psychophysiology is defined as the "study of brain-behavior relationships in the framework of peripheral and central physiological responses" (Hugdahl, 1995, p. 8). Within the human body, emotional regulation is implemented in the central nervous system (CNS), which includes the brain and spinal cord, and the peripheral nervous system (PNS), which is composed of cranial and spinal nerves, ganglia, and sensory receptors (Taylor, 2016). The peripheral nervous system is divided into the somatic nervous system, which attends to muscles that produce overt behaviors, and autonomic nervous system (ANS), which hosts the sympathetic and parasympathetic divisions. The sympathetic division provokes the body into fight or flight responses in situations of stress, danger, excitement, exercise, or emotional distress. Some examples of sympathetic responses are shortness of breath, sweaty palms, and an increased heart rate (Cannon, 1929; Gregersen et al., 2014; Turpin, 1986; Wolfe, 2006). Conversely, the parasympathetic division counteracts sympathetic reactions through calming down the body to promote activities such as rest and digest (Lang et al., 1993; Taylor, 2016). The sympathetic and parasympathetic divisions are subconscious, involuntary reactions that can be measured, making them prime candidates for understanding individuals' unconscious emotional responses at a purer, less cognitively-filtered level, which bypass possibly obscured self-reports (Surwillo, 1990). A comprehensive review of the biological origins of emotion is beyond the scope of this dissertation; however, it is critical to note that psychophysiology attempts to tap into the primary measures of emotional responses. In essence, that is the main assumption of psychophysiology, to which I will now attend.

Psychophysiology

Assumptions and caveats.

Generally speaking, *psychophysiology* refers to the study of the informative and meaningful relationship between the mind (*psych* = “soul/spirit” in Greek) and body (*physio* = “nature” in Greek). The primary assumption within psychophysiological research is called the brain-body interface. The first part of the assumption states that human perception, thought, emotion, and action are embodied and embedded phenomena; secondly, that the measures of the processes of the corporeal brain and body contain information that can shed light on the human mind (Cacioppo, Tassinary, & Berntson, 2007, p. 14). Another way of describing the link is that behavioral, cognitive, emotional, and social events all are mirrored in physiological responses, such that psychophysiological responses may be considered “windows” into the brain and mind (Hugdahl, 1995, p. 5). The brain-body interface is requisite to accepting the findings of the present study in that we assume the physiological changes observed by the emotional induction represent, at least to a certain extent, human experiences of thought and emotion (Sequeira, Hot, Silvert, & Delplanque, 2009).

There are at least three caveats that must be acknowledged regarding the measurement of physiological data. These emphasize the complexity of measuring emotion. First, researchers have stated that there are no direct one-to-one relationships between emotions and changes in autonomic activation; that is, ANS responses are not determinedly linked with specific elicited emotions (Cacioppo et al., 2007; Christie & Friedman, 2004; Kreibig, 2010; Lane, Nadel, Allen, & Kaszniak, 2000). One such reason is that various non-emotional variables, such as participant posture, room temperature, motor activity, or simultaneous cognitive tasks that require attention may also contribute to changes in ANS measures (Kreibig, 2010).

Secondly, changes in ANS measures due to emotional stimuli are time-sensitive and observable for a range of only seconds to minutes after a stimulus is presented (Ekman, 1984; Ekman & Davidson, 1994). Appropriate caution must be used therefore when interpreting ANS measures during specific phases of event-related stimuli (Cacioppo et al., 2007; Mauss & Robinson, 2009).

Finally, it remains to be determined how well physiological measures correlate to one another. Until now, findings are inconclusive and vary widely depending upon the specific measures collected, dimensions of the emotional stimuli (i.e., length, intensity, specific/general emotion targeted), and the type of analysis performed (Cacioppo et al., 2007; Christie & Friedman, 2004; Lane et al., 2000; Lazarus, Speisman, & Mordkoff, 1963; Malmstrom, Opton, & Lazarus, 1965). Some researchers have concluded that although the measures of heart rate and skin conductance illuminate physiological arousal, these measures are largely independent of one another (Croft et al., 2004; Eysenck, 2014; Lacey, B. C. & Lacey, J. I., 1974; Lazarus & Folkman, 1984; Marschark, Richman, Yuille, & Hunt, 1987; Tremayne & Barry, 2001). Revelle and Loftus (1992) addressed the issue of unrelated measures of arousal of the peripheral system when they wrote: “unfortunately, these measures of the hand, the heart, and the head do not produce evidence for a unified arousal system” (p.121). This statement was not an argument against the validity of the arousal construct, but rather an observation of the variability within measures of the same construct. One theory that could explain the apparent disorder of psychophysiology is that of directional fractionation, which is the observation that an individual’s physiological response system may be independent of, or even opposite to, the direction of change in another response system, resulting in misalignment, or divergent fractionation, of psychophysiological measures (Hugdahl, 1995, p. 45). In order to attend to the possibility of divergent psychophysiological measures, the present study will measure three physiological measures. Similar to the recommendation of collecting multiple measures of emotion, psychophysiology researchers advocate collecting multiple physiological measures in order to have as much information as possible (Croft et al., 2004; Hilton, 1975; Kreibig, 2010; Kreibig et al., 2007; Malmstrom et al., 1965; Schneiderman & McCabe, 1989).

The three measures. The following three sections offer an overview of the three physiological measures employed in the present study: heart rate, heart rate variability, and skin conductance levels.

Heart rate (HR). Heart rate is the number of heartbeats per minute (bpm). *Figure 1* is an example of a normal heartbeat readout from an echocardiogram (EKG). Average heart rates for adults fall between 60-100 beats per minute (Association, 2015).

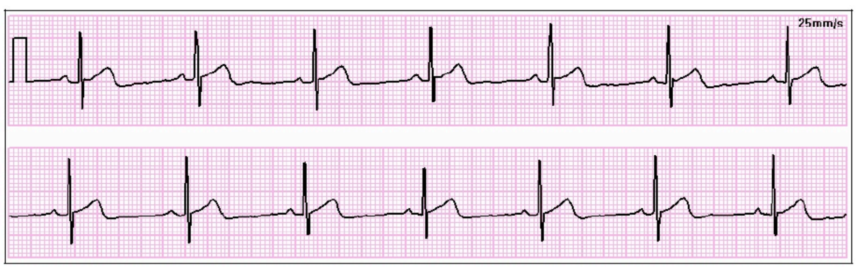

Figure 1. Sample heart rate (Klabunde, 2016)

Although widely used in the field of psychophysiology, heart rate measures do not represent a clear activation of either the parasympathetic or sympathetic nervous system. This is due to the fact that the heart is controlled by multiple internal and external mechanisms. Internally, the sinoatrial node, known as the heart's pacemaker, dispatches an electrical charge that prompts rhythmic contractions of the heart. Externally, nerves from both the ANS and CNS innervate, or supply nerves to, the heart. Specifically, the Vagus nerve is the parasympathetic representative in the heart, and it is responsible for activating the release of the neurotransmitter acetylcholine in order to decelerate the heart rate and promote “rest and digest” activities (Porges, 1995; Porges, Doussard-Roosevelt, & Maiti, 1994). Finally, reflex mechanisms controlled by the different structures in the CNS also influence heart rate: medulla, hypothalamus, cerebellum, and amygdala (Surwillo, 1990). Due to the interplay of the sympathetic and parasympathetic divisions of the ANS within the heart, researchers agree that heart rate measures do not serve as accurate indicators of either arousal or valence (Berntson, Quigley, & Lozano, 2007; Carvalho et al., 2012; Frazier, Strauss, & Steinhauer, 2004). It is for this reason that I took issue with the work spearheaded by Gregersen et al. (2014), the pioneering SLA researchers who employed HR measures as an indicator for arousal, and foreign language anxiety. Due to the dual-innervation of the heart, it was unreasonable to draw the connection between changes in heart rate and such specific emotions such as foreign language anxiety.

Despite the challenge of interpreting heart rate, mood induction experiments have found converging heart rate patterns over time. For instance, participant heart rate patterns during oral presentations generally followed a trend of initial increases in heart rate followed by an eventual deceleration to normal rates (Croft et al., 2004; Witt et al., 2006). Other examples of converging heart rate patterns come from studies in which participants were threatened with electrical shock (Elliott, 1974), exposed to photos of snakes (Klorman, 1974) or spiders (Hare, 1973), or shown scenes of mutilation (Klorman, Weissberg, & Wiesenfeld, 1977); unsurprisingly, these participants exhibited heightened heart rates, which represents their activated sympathetic nervous system.

In an attempt to discern how heart rate compares to the additional measures of physiology and self-report, and to have some aspect of comparability with other studies in this line of research, I include measures of heart rate in the present study. Additionally, I will investigate heart rate variability and skin conductance levels, which have been found to serve as better indicators of parasympathetic and sympathetic responses.

Heart rate variability (HRV). Heart rate variability is defined as the "oscillation in the interval between consecutive heart beats as well as the oscillations between consecutive instantaneous heart rates" (Citi, Brown, & Barbieri, 2012, p. 354). In other words, heart rate variability is the measure of the average variance of the distance between heartbeats (IBI, inter-beat intervals) over a particular measure of time. *Figure 2* demonstrates the basic concept of how HRV is calculated. The numbers near the peaks of the heart rate curve represent the IBIs, from which an average HRV level is calculated over a determined amount of time.

Figure 2. Sample heart rate variability (Marker)

HRV is predominantly understood to be a trait level measure of one's parasympathetic nervous system; that is, an individual's cognitive, emotional, and autonomic self-regulatory capacity in the face of stressful or demanding situations (Beauchaine, Gatzke-Kopp, & Mead, 2007; Delaney, 2012; Frazier et al., 2004; Thayer, Hansen, Saus-Rose, & Johnsen, 2009; Thayer & Lane, 2009). As previously mentioned, the parasympathetic system calms the body during and after stimulating events; thus, higher levels of heart rate variability indicate better parasympathetic ability to regulate emotions. This metric has gained popularity among professional athletes and practicing mental health care providers as a mechanism of cardiorespiratory biofeedback training for strengthening homeostasis in the baroreceptors (Lehrer & Gevirtz, 2014; Martin, 2012; Michel, 2016).

The critical contributor in calming down the heart is the Vagus nerve, and *vagal tone* is a term used to describe the responsiveness of the parasympathetic system (Grossman & Taylor, 2007; Porges, 1995; Porges et al., 1994). High levels of HRV have been found to be associated with better self-regulatory ability, adaptability, attention control, coping strategies, positive emotion, social well-being, and the ability to make faster and more accurate decisions in cognitive tasks requiring executive function (Brosschot, Van Dijk, & Thayer, 2007; Fabes & Eisenberg, 1997; Hansen, Johnsen, & Thayer, 2003; Kok & Fredrickson, 2010; Oveis et al., 2009; Park, Van Bavel, Vasey, & Thayer, 2012; Thayer et al., 2009; Thayer & Lane, 2009). Correspondingly, studies have shown that those with low levels of HRV have poor self-regulatory capacity, a reduced ability to adapt their emotions to changing situations or expectations, and more frequently observed rigid and hyper-vigilant behavior (Gross, 1999; Gross & Muñoz, 1995).

While many researchers consider HRV to be exclusively a trait level measure, there is increasing evidence that it may also measure state level changes (Park, Vasey, Van Bavel, & Thayer, 2013; Thayer, Friedman, & Borkovec, 1996). Correlations reveal a strong relationship between high tonic HRV and phasic HRV; thus, phasic HRV may be considered a measure of state changes in the parasympathetic activity of emotional regulation and protective functioning toward threat (Beauchaine et al., 2007; Dawson, Schell, & Filion, 2007; El-Sheikh, Hinnant, & Erath, 2011). For instance, previous research involving the induction of mental stress through arithmetic calculation resulted in decreased levels of high-frequency HRV (Pagani et al., 1989; Pagani et al., 1991).

In the present study, I will use HRV to represent parasympathetic ability of participants and how well they regulate potential stress from the film stimuli. I will measure baseline measures to serve as trait-level metrics and active phase measures as state-level. This in itself will be interesting to observe whether HRV exhibits changes over a short amount of time and thus can offer evidence toward whether HRV is a feasible measure of state-level changes.

Skin conductance levels (SCL). Among all psychophysiological measures, SCL is considered to be the most sensitive index of arousal (Barry & Sokolov, 1993; Croft et al., 2004; Tremayne & Barry, 2001). It is also one of the most widely used measures because it is noninvasive, easy to collect, costs relatively little, and has a straightforward interpretation of data (Cacioppo & Tassinary, 1990; Cacioppo et al., 2007; Carvalho et al., 2012). Research from a variety of fields, ranging from advertising (Bolls et al., 2001; Poels & Dewitte, 2006) and game theory (Cowley et al., 2014; Cowley et al., 2016) to psychology employ this method to tap into constructs such as vigilance, perception, decision-making, lie detection, motivation, and emotion (Amiez, Procyk, Honore, Sequeira, & Joseph, 2003; Sequeira et al., 2009).

Skin conductance level (SCL) is the measure of time an electrical current travels from one electrode to another. Two electrodes are placed on two adjacent fingers on the non-dominant hand (Fowles et al., 1981; Lykken & Venables, 1971), and when the ANS is activated, liquid in the eccrine glands (many of which are located in the palm of hands and soles of feet) builds up, crosses a threshold level (e.g., the sweat barrier) of the outermost layer of skin, and appears on the surface of the skin. This liquid, known as sweat, accelerates the conductance of the electrical current between the two electrodes (Boucsein, 1992; Dawson et al., 2007; Dawson, Schell, & Courtney, 2011). Sweat transpires for purposes of both thermoregulation and behavioral or psychological functions. The location of the sweat glands dictates the type of function it is serving; sweat glands found in palms and fingers are posited to be most responsive to sensory stimuli and psychological arousal rather than changes in body temperature (Edelberg, 1972; Surwillo, 1990). *Figure 3* demonstrates an event-related response of the electrodermal (EDA) system. At the left of the figure, there is a stimulus that activates the rise in skin conductance levels after a brief moment of latency. Following the peak of skin conductance response, the SCL gradually declines back to resting, or tonic, levels.

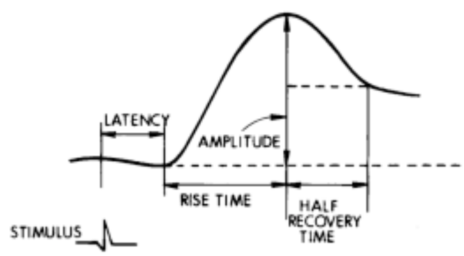

Figure 3. Skin Conductance Event Related Response (Dawson et al., 2007)

Key terms

The following key terms will be employed throughout the remainder of the present study. *Trait*, or *tonic*, is a term used to describe emotions or behaviors that are of a more permanent or consistent nature over time and situation; conversely, *state*, or *phasic*, emotions or behaviors are the transient changes due to an altered condition. These terms were briefly introduced in Chapter 1 in relation to anxiety. For example, Spielberger (2010) developed the State-Trait anxiety inventory, which attempted to discriminate between in the moment (state-level) and a general tendency (trait-level) for anxiousness.

The following set of terms originates from literature concerning stimulus-response studies. In brief, a *stimulus* is an external agent that may cause an individual to attend more closely to the novel object (*appraisal*) and respond with different behaviors (*orienting* or *defense*). In the present study, I use the term *stimulus*, or *stimuli*, to refer to the emotion-eliciting film clips. *Appraisal*, or cognitive evaluation, is a term that describes an individual's initial response and attendance to unexpected or novel stimuli (Christianson, 2014). This term is closely related to the two behavioral responses of appraisal, *orienting* and *defense* (Hackley, Boelhouwer, Lang, Simons, & Balaban, 1997). *Orienting* is considered to be the "gateway to attention" in that this response increases an individual's attendance to stimuli to allow for maximal input (Cook & Turpin, 1997; Hackley et al., 1997, p. xix; Lacey, J. I. & Lacey, B. C., 1974). In general, the physiological indicator most associated with orienting is a deceleration of heart rate throughout the exposure of the stimulus. The *defense* response, on the other hand, is characterized by an individual's reduced capacity for input absorption and desire to remove oneself or escape from the threat (Lang, Simons, & Balaban, 2013; Sokolov, 1963). The predominant physiological response of the defense response is a heightened activation of the sympathetic nervous system, such as an elevated heart rate or skin conductance level. Finally, Dawson et al. (2007) defined *habituation* as the "ubiquitous and adaptive phenomenon whereby subjects become less responsive to familiar and non-significant stimuli" (p. 167). This occurrence can be observed through decreases in amplitude and reduced response activity in particular psychophysiological measures (Barry & Sokolov, 1993).

Summary

In this chapter, I introduced the constructs of anxiety and emotion from the standpoint of psychology, offered a brief review of the elicitation and measurement, and explained underlying assumptions and considerations of psychophysiology. What is important to note is the apparent lack of psychological literature, barring the subfield of bilingualism, regarding issues of SLA within cognitive psychology and cognitive science. This is puzzling due to the fact that both fields investigate similar constructs such as attention and cognition, sensory input, and individual differences (Tomlin & Villa, 1994).

The present study

Before I introduce the purpose and procedure for the present study, I would like to review how SLA and psychology serve as the two towers that support the forthcoming argument (cable), hangers (operationalizations), and deck (analysis and results of the present study) of the bridge-building enterprise. In chapters 1 and 2, I reviewed relevant research and introduced assumptions for the present study. Most notable is the understanding that the commonly known foreign language anxiety represents just one of many types, durations, and intensities of anxiety, and emotion more generally. From the perspective of the tower of psychology, we gained a more dimensional understanding of emotion and various methods of elicitation and measurement. The present study spotlights the psychophysiological dimension of emotion and how it relates to a subsequent vocabulary-learning task, which is the operationalization of intentional learning and cognition.

Extant research at the intersection of psychophysiology and SLA is sparse (Chwilla et al., 2011; Gregersen et al., 2014; Verhees et al., 2015). As previously mentioned, the most closely related study was Gregerson et al. (2014). This study provided a starting point for a more methodologically rigorous investigation of the relationship between emotions, physiology, and language. Two areas of concern rise to the surface. First, heart rate was reported as an indicator of both sympathetic and parasympathetic responses of the ANS, and thus, an unreliable indicator for either valence or arousal (Croft et al., 2004; Witt et al., 2006). Secondly, the use of heart rate alongside other measures of foreign language anxiety suggests that it represents the direct physiological manifestation of anxiety, a discrete emotion, and also the more specific foreign language anxiety. Such a relationship can only be determined through methodologically standardized procedures and analyses of multiple physiological measures that more closely discriminate specific anxious responses.

Research questions. The aim of the present study is to add one more brick onto the bridge that spans the interdisciplinary gap between SLA and psychology. First, I will use psychophysiological and self-report measures to operationalize emotions. Second, I will employ a mood induction paradigm to determine how externally manipulated emotions influence vocabulary learning. Finally, I will extend the scope of emotion research beyond the inquiry of foreign language anxiety by introducing a broader, more psychologically grounded operationalization of emotion. The overarching research question of the present study is how emotions—represented by self-report and psychophysiological responses—influence cognition—represented by a paired-associates vocabulary-learning task in a novel language.

The present study is motivated by the following research questions:

1. What self-reported and physiology measures show clearest changes under the influence of the film-based mood induction?
2. How does a film-based mood induction influence vocabulary learning on a paired-associates learning task?
3. What are the effects of physiology on vocabulary learning on a paired-associates learning task (a) in isolation, and (b) when controlling for the mood induction, self-reported measures of state-level emotional changes, and self-reported measures of trait-level anxiety?

Hypotheses. For research question 1, I hypothesize that the emotional manipulation check for the self-report portion of the study will reveal significant differences between the negative and neutral mood induction group; that is, that the negative group will report lower levels of valence (happiness) and higher levels of arousal (calmness) following the film clips (Fox, 2016; Fox & Miller, 2015; Miller et al., 2017). As for physiological changes during mood induction, I expect to observe group level differences: the negative group will exhibit HR deceleration (Carvalho et al., 2012; Codispoti et al., 2008), decreases in HRV (Demaree et al., 2004; Demaree, Robinson, Pu, & Allen, 2006; Frazier et al., 2004; Kreibig et al., 2007; Pauls & Stemmler, 2003), and increases in SCL (Cacioppo & Tassinary, 1990; Carvalho et al., 2012; Codispoti et al., 2008; Kreibig et al., 2007; Palomba et al., 2000). I anticipate the neutral mood induction group to exhibit either no change whatsoever and/or a combination of the following changes: HR deceleration, and decreases in HRV and SCL (Kreibig et al., 2007).

As for research question 2, I expect that negative mood induction will have an overall detrimental effect on all vocabulary posttests (Fox, 2016; Fox & Miller, 2015; Miller et al., 2017). Finally, for research question 3 part A, I hypothesize that, of the three physiology variables entered into the physiology-only analyses, SCL, followed by the HRV measure, will be most informative in predicting vocabulary learning. As for part B, I expect that both physiology and self-report measures will influence vocabulary learning; particularly those participants with high trait levels of HRV and SCL will perform better (Dawson et al., 2007; Frazier et al., 2004; Huang-Pollock, Carr, & Nigg, 2002; Jönsson, 2007; Thayer et al., 2009; Thayer & Lane, 2009).

At this point in the bridge construction, I have established the relevant information from the two towers and have extended the cable–the argument and hypotheses–between them. In the forthcoming chapters I outline the materials and methodologies I will employ to operationalize the bridge construction (chapter 3), report the results of the project (chapter 4), and discuss the results in light of the research questions (chapter 5).

CHAPTER 3

Methodology

Participants

The participants in this study were 70 right-handed individuals (18 males, 52 females) from a large Midwestern university in the United States. The reason we required right-handed individuals take part in the study was because the skin conductance electrodes were placed on the left-handed fingers of the participants. Portions of the experiment required the participants' ability to record responses using a pen, so it was required that their right hands be the dominant hand. The participants' mean age was 23.5 (SD = 6.57). I conducted telephone screenings to ensure that all participants met the inclusion criteria: they (a) were over 18 years old; (b) had no previous experience with the Arabic, Indonesian, and Malaysian languages; (c) were willing to complete the two-part study; and (d) were aware that they may be exposed to potentially graphic or violent film clips. Participants were randomly assigned to a comparison group (n = 35) or a negative emotional group (n = 35). Less than five participants stated that they were sensitive to potentially graphic or violent film clips, and they were placed into the neutral mood induction group. The gender of the participants was equally distributed between the two groups. Each participant gave written informed consent, was free to withdraw from the study without penalty, and was compensated with $25.

Materials

As previously mentioned, this study is part of a larger research project involving multiple measures and research questions (Fox, 2016; Fox & Miller, 2015; Miller et al., 2017). The instruments listed below are those that were analyzed in the context of this dissertation:

emotional stimuli, psychophysiological measurement instruments, emotional self-reports, a paired-associates vocabulary-learning task and posttests, and anxiety questionnaires.

Emotional stimuli. I utilized movies from an established affective movie database (referred to as the Emotional Movie Database, or EMDB, [Carvalho, et al., 2012]) to induce changes in participants' valence and arousal. As a reminder, valence is the degree of happiness or sadness, and arousal represents the level of anxiousness or relaxedness that is reported by participants on a 9-point Likert scale. All film clips were soundless, 40 seconds in length, and maintained a resolution of 720 x 576. Two neutral film clips were also selected for pre-mood-induction practice for both groups. Following practice, the negative emotional group viewed six horror film clips from the EMDB, which elicit low arousal and mid-level valence ratings known to elicit high arousal and negative valence in self-reports. The neutral emotional group viewed six films from the scenery genre of the EMDB. Table 1 displays the film clip descriptions and mean ratings for valence and arousal for the practice, neutral, and negative films, respectively, reported in a recent study (Carvalho et al., 2012).

Table 1

Mean ratings for valence and arousal for films

Clip Type	*Clip No.*	*Name*	*Description*	*Valence Mean (SD)**	*Arousal Mean (SD)**	*SCL Variation Mean (µs)***	*HR Variation Mean (bpm)***
Practice	6000	Homemade footage 1	Moving objects on a table	4.90 (1.82)	2.44 (2.23)	-1.77	-3.271
Practice	6001	Homemade footage 2	Moving objects on a table	4.74 (1.51)	2.33 (1.97)	-1.77	-3.271
Neutral	5000	Disney's Earth	Desert and polar scenes	5.88 (1.99)	2.99 (2.25)	-1.813	-4.460
Neutral	5001	Disney's Earth	Mountains with ice	5.83 (1.67)	2.72 (2.03)	-1.813	-4.460
Neutral	5002	Disney's Earth	Scenery with polar scenes and the dusk	5.68 (1.70)	2.51 (1.86)	-1.813	-4.460
Neutral	5005	Disney's Earth	Sandstorm and desert	5.32 (1.06)	2.86 (1.97)	-1.813	-4.460
Neutral	5008	Disney's Earth	Several scenes from a jungle and in the end mushrooms start to grow	5.73 (1.54)	2.54 (1.95)	-1.813	-4.460
Neutral	5009	Disney's Earth	Clouds swirling	6.17 (1.44)	2.79 (1.80)	-1.813	-4.460
Negative	1000	The Ruins	Amputation scene on top of the ruins	2.04 (1.98)	7.11 (1.77)	.742	-7.514
Negative	1005	Midnight Meat Train	Leslie Bibb inside a carriage with bodies hanging from the ceiling	2.06 (1.48)	6.92 (1.74)	.742	-7.514
Negative	1008	The Rest Stop	Jaimie Alexander giving a merciful shot to the head of the police	1.94 (1.41)	6.53 (2.05)	.742	-7.514
Negative	1009	Midnight Meat Train	Vinnie Jones with a vicious attack direct to a woman that ends with the decapitation of the victim	1.83 (1.24)	6.88 (1.70)	.742	-7.514

Self-Assessment Manikin. Emotional self-reports in the form of self-assessment manikins (SAMs) (Bradley & Lang, 1994) were administered throughout the course of the experiment. Participants marked their current levels of valence, arousal, and dominance on a nine-point Likert scale. This scale has been widely used with other psychological research as a manipulation check of the emotional stimuli (Kreibig, 2010; Oveis et al., 2009; Volokhov & Demaree, 2010). Previous research has employed this scale during mood induction paradigms using film clips to verify that participants who viewed negative film clips reported the expected decreases in valence levels (i.e., became less happy) and expected increases in arousal levels (i.e., became more anxious) (Carvalho et al., 2012; Kreibig et al., 2007). See Appendix A for the measure.

As previously discussed in chapter 2, the sole use of self-report to validate an experimental mood induction is not recommended. It is entirely possible that a participant does not accurately acknowledge his or her mood shifts or their intensity. For that reason, I include physiological measures that correspond to emotional responses alongside the self-report data to examine how self-report corresponds or diverges from this unconscious, and thus purer, measure (Cowley et al., 2014; Cowley et al., 2016).

Vocabulary training stimuli. I employed a paired-associate learning task using 24 Indonesian words for learning. I chose Indonesian as the target language because it is a less-commonly-taught language (Lagere, 2015) and uses a Romanized script. This decreased the likelihood that participants would have been previously exposed to it, while ensuring that the orthography was accessible to them. The vocabulary items were concrete disyllabic nouns of varying lengths (8 four-lettered, 8 five-lettered, and 8 six-lettered) that were not English cognates (de Groot & Smedinga, 2014). Examples include *batu* (rock), *jalan* (road) and *kertas* (paper). Appendix B contains the 24 target words, plus 3 practice words, with matching illustrations from the International Picture Naming Project (Szekely et al., 2004). A sample of the lexical training is displayed in Appendix C.

Vocabulary posttests. I administered three posttests immediately after the vocabulary training phase: (a) a free recall activity eliciting any English and Indonesian words seen during training, (b) a productive word recall task eliciting the corresponding Indonesian words for all 24 English labels, and (c) a receptive word recall task eliciting the English translation of the Indonesian words. I randomized the presentation order within each test so that participants would not detect patterns. Each participant had 2 minutes 24 seconds for each of the posttests. For procedures and scoring the three vocabulary posttests, I followed the approach in a similar study developed by Barcroft (2003). The inter-rater reliability for the two scorers was high (α = .99) for each of the posttests. I, along with my research partner, scored all tests using the Lexical Production Scoring Protocol adapted from Barcroft (2003). Inter-rater reliability for all six posttests was high (α >.99), which allowed me to average the two ratings for participants' absolute scores. See Appendix D for the posttest instructions and Appendix E for the scoring rubric.

Anxiety questionnaires.

Foreign Language Classroom Anxiety Scale (FLCAS). (Horwitz et al., 1986). The questionnaire is composed of 33 statements presented with a five-point Likert scale (1 = strongly disagree, 5 = strongly agree) that reveal attitudes regarding participants' study of a foreign language. The range of possible scores on the measure is 33-165 points. An exploratory factor analysis of the FLCAS questionnaire reported in Fox et al. (in preparation) yielded two factors that accounted for nearly 50% of the variance: foreign language anxiety and learning anxiety (Castaneda, Palermo, & McCandless, 1956; Cubukcu, 2008; MacIntyre, 2002; Oxford, 2014; Schein, 2002). See Appendix F for the questionnaire.

Mood and Anxiety Symptom Questionnaire (MASQ). (Seidenberg et al., 1994; Watson & Clark, 1991). Commonly administered in the field of psychology, this questionnaire attempts to differentiate two specific aspects of anxiety: anxious arousal (17 items) and anhedonic depression (22 items), which is the loss of interest in activities that once were enjoyable to the individual. This 38-item questionnaire utilizes a five-point Likert scale (1 = not at all, 5 = extremely) on which participants must respond to a statement. The range of possible scores is 38-195. In the study by Fox (2016); Fox et al. (in preparation), two factors emerged from the MASQ and accounted for 37% of the variance: positive emotionality and somatic distress (Seidenberg et al., 1994). See Appendix G.

Penn State Worry Questionnaire (PSWQ). (Fresco et al., 2002; Meyer, Miller, Metzger, & Borkovec, 1990). This questionnaire targets pathological worry, defined as "a chain of thoughts and images, negatively affect-laden and relatively uncontrollable; it represents an attempt to engage in mental problem-solving on an issue whose outcome is uncertain but contains the possibility of one or more negative outcomes; consequently, worry relates closely to the fear process" (Fresco et al., 2002, p. 314). The questionnaire is composed of 16 five-point Likert scale items (1=not at all typical of me, 5=very typical of me). Consistent with previous factor analyses (Fresco et al., 2002), Fox et al. (in preparation) found one factor, worry, which accounted for 51% of the variance. See Appendix H.

State-Trait Inventory for Cognitive and Somatic Anxiety—Trait (STICSA). (Gros, Antony, Simms, & McCabe, 2007; Spielberger, 2010). This part of a two-set questionnaire taps into one's self-reported levels of anxiety as a long-term, permanent trait. This contrasts with *state* anxiety, which is considered to be more transient. This questionnaire is composed of 21 items on a four-point Likert scale (1 = not at all, 4 = very much so). Within psychological literature, the

STICSA-Trait questionnaire has consistently produced two factors of trait anxiety, cognitive symptoms of anxiety (10 items) and somatic symptoms of anxiety (11 items). Fox (2016) also performed an EFA that produced two underlying factors consistent with previous research: cognitive distress and somatic distress (Gros et al., 2007). See Appendix I.

Experimental task and procedure

The experiment took place over the course of two days. Figure 4 represents the procedure graphically.

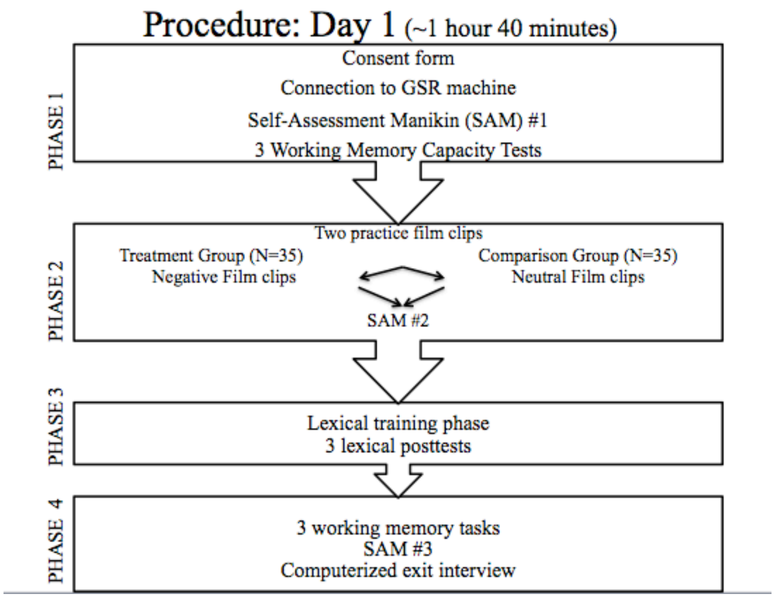

Figure 4. Day 1 procedure

Day 1.

Phase 1: Preparation. Participants from all groups performed Phase 1 in the same manner. Participants sat in front of a computer monitor, were connected to the BioSemi Active 2 system for psychophysiological data collection, and completed both an initial emotional self-report and a single block of three complex WMC tests (operation span, symmetry span, and non-word repetition [Baddeley, 1992, 2003; Baddeley et al., 1998; Foster et al., 2015]) to establish a working memory capacity baseline measure.

Phase 2: Emotional induction. First, all participants viewed two practice clips considered emotionally neutral movie clips in order to provide baseline heart rate and skin conductance measures. After the practice session, the mood induction began and participants in both groups viewed one 40-second film clip of either neutral or negative content followed by 10 seconds of a blank screen to allow for a return to baseline. Participants viewed six film clips in this manner. An immediate emotional self-report was conducted upon the completion of the sixth film clip. *Figure 5* displays the experimental timeline of the emotional stimuli presentation; this figure is important to understand for later analyses (cf. Extraction of psychophysiological data section).

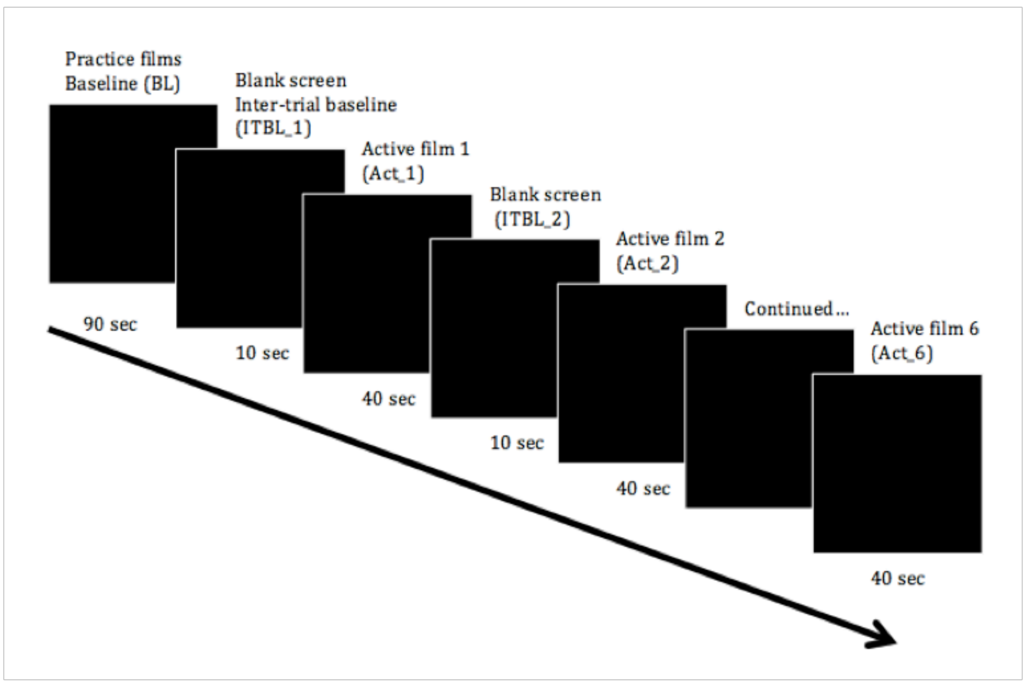

Figure 5. Experimental timeline and measures

Phase 3: Vocabulary. After the mood induction, participants from each group engaged in paired-associates vocabulary learning similarly. Participants remained in front of the computer and saw a target word (left-hand side of the screen, 64-point, Arial font), accompanied by the English translation to the right and its pictorial representation above the word pair for 8 seconds per trial. One complete block included the presentation of all 24 trials in this manner, and participants viewed three blocks in total, each in a randomized order to reduce pattern development. Thus, participants studied each target item for a total of 24 seconds, which according to Barcroft (2003) provides enough time to avoid ceiling and floor effects of learning during the training session. Upon completion of the vocabulary-training phase, all participants completed three immediate vocabulary posttests.

Phase 4: Completion. Following the posttests, participants completed all three working memory capacity tests again, an emotional self-report, and a computerized exit interview. The rationale for repeating the working memory capacity tests was to determine whether the mood induction impacted working memory capacity in the participants (for results, see Miller et al., 2017). See *Figure 6*.

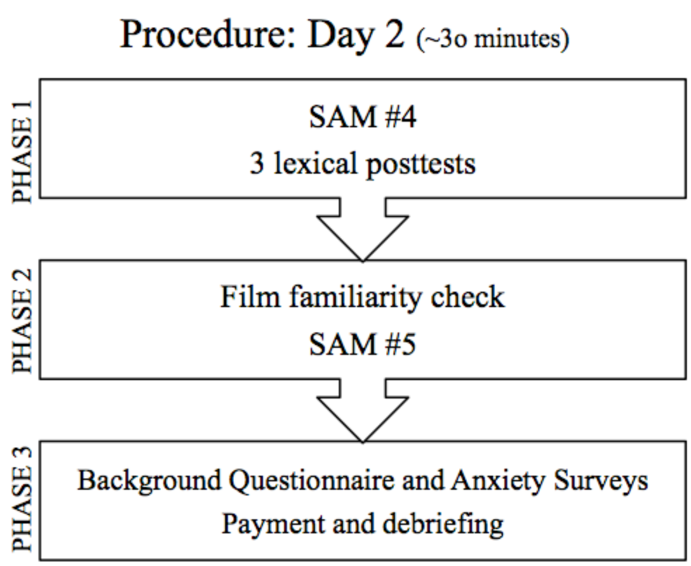

Figure 6. Day 2 procedure

Day 2. Day two of the experiment occurred 48 hours after day 1 and included 4 phases. The figure above displays the procedure for the second day of the experiment.

Phase 1: Vocabulary tests. Participants completed an initial emotional self-report and all three delayed vocabulary posttests. The same posttest format and time restrictions as in Day 1 were followed. For each delayed posttest, word presentation orders were randomized from their corresponding initial posttests. Similar to Barcroft (2003), participants were asked to indicate if they had discussed or studied the experimental words between the posttests.

Phase 2: Film check. Participants viewed 5-second clips of the same film clips as Day 1 and responded "yes," "no," or "I don't know" on a form asking if they had seen these films before the experiment. At the conclusion of viewing the truncated film clips and completing the form, participants completed a final emotional self-report.

Phase 3: Debriefing. Participants completed the four anxiety questionnaires and a background questionnaire including information about their gender, age, previous exposure to Indonesian, Malaysian, and Arabic languages, and film preferences.

Psychophysiological data

Extraction. Participants' electrocardiogram (ECG) and electrodermal activity (EDA) readings were collected using the BioSemi Active 2 system (BioSemi, 2011), a psychophysiological data collection system, sampling at 1000 Hz. The data were then saved on a computer running the Brain Vision Analyzer software (Analyzer). Although psychophysiological measures may be appear to have the most objectivity in the study of emotion, there is widespread lack of consensus in the field regarding which specific procedures to standardize. This is despite the fact that psychophysiologists have published recommendations for standardizing the extraction, measurement, and reporting procedures (Berntson et al., 1997; Boucsein et al., 2012; Cacioppo et al., 2007; Cardiology, 1996; Cowley et al., 2016; Fowles et al., 1981; Mauss & Robinson, 2009).

As previously shown in the *Figure 5*, the following three physiological measures were extracted over the same time segments for each participant based upon the hand-recorded time segment of the beginning and end of each film clip: baseline segment (1 x 90sec), six active phase segments (6 x 40sec), and six inter-trial baseline segments (6 x 10sec) used for baseline-correction.

Each physiological measure was extracted following recommendations of previous research, and continuous data points representing HR, HRV, and SCL were used in subsequent analyses. These measures are operationalized in the following manner: first, an overall baseline measure, is the mean HR, HRV, or SCL during the 90 seconds of the practice film clips. Second, the active state HR, HRV, or SCL is the baseline-corrected mean score during the six 40-second film clips (Cardiology, 1996; Carvalho et al., 2012; Frazier et al., 2004; Oveis et al., 2009; Park et al., 2013; Volokhov & Demaree, 2010). Below I outline the specific cleaning, processing, and scoring procedures for each measure.

Scoring

Heart rate. Following Carvalho et al. (2012), heart rate (measured in beats per minute; BPM) was obtained using a 3 lead ECG, with a lead II configuration. Two electrodes were placed on the left side of the body between the ninth and tenth rib, and at mid-sternum. Data was read into the Brain Vision Analyzer software program (Analyzer, 2006) in order to isolate the heart rate by subtracting one lead from the other to reduce the effect of noise in the data collection. Initially, nineteen of the 70 (27%) participants' complete heart rate data were unreadable by the Brain Vision Analyzer software, so their data were not analyzed for either the heart rate or heart rate variability because HRV employs the same heart rate data that is inputted into a different software program described below. Such a large loss of physiological data due to technological issues or individual differences is expected in this field of study (Tran et al., 2007). Nonetheless, 51 participants' data were available for further investigation.

Each participant had 13 segments of data to inspect: 1 90-sec baseline measure, 6 10-sec inter-trial baseline measures, and 6 40-sec active film phase measures. All of the procedures outlined below were performed on each of the participant's 13 segments. First, using the Brain Vision Analyzer software, I visually inspected the R-waves of the electrocardiogram for artifacts, which are errors in the computer's algorithmic reading of the data (e.g., deviant inter-beat intervals (IBIs)). If artifacts were found, I marked the troublesome section of the R-wave and later applied an interpolation [standardized adjustment calculation] function to these sections. Of the 51 participants' data, there were 109 time segments (16%) segments in that required interpolation. Next, the files were exported from BVA as a text file. I then calculated inter-beat intervals (IBI), the time between the peaks of adjacent heartbeats, and heart rate per minute (IBI divided by 60000 = beats per minute, BPM) for each second of the segment. The heart rate per minute was then mean aggregated for each time segment.

For the 109 time segments that required interpolation, I employed the weighted averages interpolation correction method, which takes the average of a consistent number of normal IBIs immediately preceding and following the error zone. For cases that were interpolated, the mean heart rate was calculated for both the raw (excluding the artifact errors) and interpolated IBIs. The correlation coefficient for these raw versus interpolated scores was .956, which indicates that the interpolation method was reliable.

Heart rate variability. Using the same inter-beat intervals (IBIs) extracted for the HR measurements, the identical raw/interpolated IBIs from the heart rate measures were read into Kubios, a heart rate variability analysis software program (Mateo & Laguna, 2003; Peltola, 2007; Tarvainen & Niskanen, 2012). As previously mentioned, the 16% of segments that had artifacts were calculated based upon the raw IBIs with an automatic medium-level interpolation selected within Kubios. All segments underwent a first order polynomial (linear) detrending filter in order to mitigate the common challenge of stationarity of the data, which is due to baseline shifts and errors caused by participant movements (Peltola, 2007; Tarvainen & Niskanen, 2012; Volokhov & Demaree, 2010).

As previously mentioned, there are multiple extraction methods to employ for HRV, and due to the relatively recent use of this metric, standardized procedures have yet to be announced (Berntson et al., 1997; Cardiology, 1996). No consensus has been reached regarding standardizing the method of heart rate variability analysis, specifically between the use of Fast Fourier Transformation (FFT) and Autoregressive (AR) transformation techniques. Ultimately, I chose the most commonly used FFT approach because this is the non-parametric method of analyzing HRV. Furthermore, (Hayano et al., 1991) reported that the two methods are highly correlated ($r = .96$). FFT analysis yields a power estimate distribution across frequencies. Power, in this context, is used to refer to the high-frequency band of the represented frequencies, which represents the quantitative estimate of vagal control of the heart. I extracted high-frequency HRV power scores in order to quantify participants' vagal control of the heart, or rather, the strength of their parasympathetic nervous system to regulate emotions (Park, Vasey, Van Bavel, & Thayer, 2014; Thayer, Hansen, & Johnsen, 2010). The units for spectral estimates of high-frequency power are milliseconds squared per Hertz ($msec^2/Hz$) within the respiratory range of 0.15 – 0.40 Hz (Berntson et al., 2007).

Figure 7 displays the frequency distribution of heart rate variability for a given individual over a certain amount of time. Along the x-axis are the frequency ranges, from very low (VLF, 0 – 0.04 Hz) and low (LF, 0.04 – 0.15 Hz), to high frequency (HF, 0.15 – 0.4 Hz). The y-axis of the graph represents the power spectral density of an individual's heart rate variability within the specified frequency ranges. As noted above, I am extracting only the high-frequency range of HRV measures.

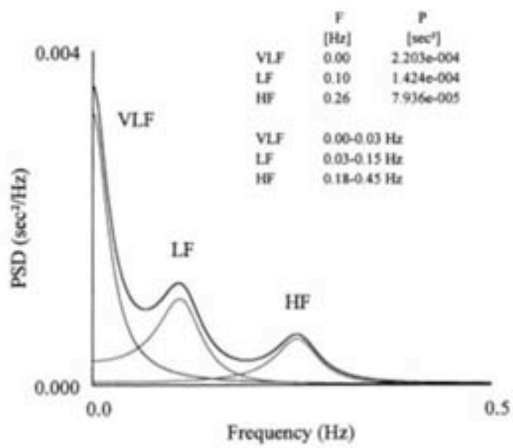

Figure Legend:

Plot of the power spectral density (PSD) computed from a stationary RR interval series from a normal subject. The power spectral density is automatically decomposed in bell-shaped curves in correspondence with the major spectral components; each component is quantified by its center frequency (F) and by its area, that is, its power (P). Power spectral density units are in s^2/Hz, whereas the frequencies are expressed in Hz. Thus, the power will be expressed in s^2. LF = 0.03 to 0.15 Hz; HF = 0.18 to 0.45 Hz; VLF = 0 to 0.03 Hz.

Figure 7. Power spectrum range of HRV (Watson, 2015)

It is important to note again that HRV measures have been primarily considered a trait-level, or tonic, measure for vagal influences of the heart. In particular, most researchers assert that accurate measures of high-frequency HRV can be gathered from a minimum amount of 1 minute of continuous data collection (Citi et al., 2012). I will inspect the phasic changes of HRV over time in the manipulation check. Based upon those findings, I may only proceed with using the overall baseline measure (from the 90 seconds of practice film clips) to operationalize this trait-level physiological measure (Citi et al., 2012).

Skin conductance levels. SCL (microSiemens; μS) was obtained with a voltage of 0.5 V from two Ag–Ag–Cl electrodes attached to a conductance module. Electrodes were attached to the second and third finger of the left hand using a conductance ointment and medical tape. The non-dominant hand was selected so that participants were still able to perform experimental tasks with the dominant right hand (Carvalho et al., 2012; Frazier et al., 2004). *Figure 8* demonstrates the placement location for the two electrodes used to measure skin conductance levels.

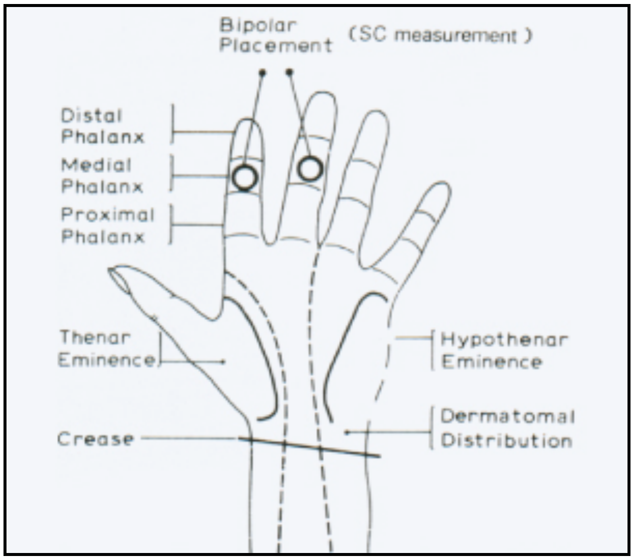

Figure 8. Skin conductance electrode placement ("Skin Conductance Explained," 2017)

In order to analyze the SCL data, I imported the data into Brain Vision Analyzer (Analyzer, 2006) and segmented it into overall baseline (90-sec), inter-trial baseline (6 x 10-sec) and active phase (6 x 40-sec) timestamps, identical to the time segments of the HR and HRV measures. These segments were then imported into MATLAB (*MATLAB User's Guide*, 1998) and analyzed offline using the Ledalab program, a plug-in specifically tailored for electrodermal activity (EDA) analysis (Karenbach, 2005). Sixty-eight of the 70 (97%) participants' skin conductance level data were unreadable by the Ledalab software. Next, the data were down-sampled to 40 Hertz and underwent a Continuous Decomposition Analysis (Benedek & Kaernbach, 2010; Cowley et al., 2014; Cowley et al., 2016). This relatively new analysis of the electrodermal signal separates the signal into tonic and phasic portions for a more accurate understanding of how the stimuli affect tonic and phasic measures independently and jointly (Benedek & Kaernbach, 2010). The measure that will be used to operationalize SCL in the present dissertation is the integrated skin conductance response ISCR for each baseline-corrected segment. The ISCR, measured in micro-Siemens (μS), improves traditional trough-to-peak or sum of the peak amplitude measures in that it employs a straightforward scoring method, avoids biases, and circumvents previously problematic issues pertaining to the ever-changing relative baselines of this measure. Benedek and Kaernbach (2010) advocate for this method of scoring the response magnitude in skin conductance data for future studies.

Statistical Analyses

In this section, I will explain the two separate factor analyses I performed to reduce the data, followed by the specific analyses for the three research questions.

Factor analyses.

Anxiety factor scores. The anxiety factor scores are operationalized as the trait-level, self-reported levels of anxiety. In order to reduce the data and determine underlying constructs, I conducted an exploratory factor analyses (EFA) on the data from each of the anxiety questionnaires (DiStefano & Mindrila, 2009). I analyzed the scores from each questionnaire by a principal axis factor analysis (PAF) with a Varimax rotation (Fabrigar & Wegener, 2011; Field, 2009; Pett et al., 2003). All Kaiser-Meyer-Olkin (KMO) values were above .7, except for the MASQ (.61, considered mediocre by Field [2013]), which confirmed the appropriateness of a factor analysis for this sample (Loewen & Gonulal, 2015). Additionally, all values of Barlett's test of sphericity were significant, and all initial mean communalities for the EFAs were above the suggested .5 benchmark (Field, 2013; Pett et al., 2003). Table 2 presents the factor analysis statistics for each initial individual questionnaire.

Table 2

Factor analysis statistics for initial individual questionnaires

Subscale Name (n items)	*Mean z-score (SD)*	*Min.*	*Max.*	*Initial Eigenvalue*	*Cumulative % of Variance*	*KMO*	*Barlett's Test of Sphericity*	*Sig.*	*Reliability Cronbach's α*
FLCAS Language anxiety (15)	0 (1)	-1.79	2.22	7.84	52.24	.90	χ^2 (105) = 640.06	< .001	.93
FLCAS Engagement (9)	0 (1)	-1.93	2.19	4.99	55.47	.84	χ^2 (36) = 351.13	< .001	.90
MASQ Positive emotionality (15)	0 (1)	-2.02	2.05	8.05	53.65	.89	χ^2 (105) = 704.58	< .001	.92
MASQ Somatic distress (4)	0 (1)	-0.76	3.84	2.42	60.56	.74	χ^2 (6) = 78.75	.001	.78
PSWQ Worry (13)	0 (1)	-1.89	1.67	7.50	57.68	.92	χ^2 (78) = 657.30	< .001	.93
STICSA Cognitive Distress (7)	0 (1)	-1.48	2.0	3.86	55.10	.81	χ^2 (21) = 210.86	< .001	.86
STICSA Somatic Distress (5)	0 (1)	-0.99	2.71	2.65	52.95	.81	χ^2 (10) = 82.24	< .001	.77

Note: FLCAS=Foreign Language Classroom Anxiety Scale, MASQ= Mood and Anxiety Symptom Questionnaire, PSWQ=Penn State Worry Questionnaire, STICSA=State-Trait Inventory for Cognitive and Somatic Anxiety—Trait

The exploratory factor analysis of the FLCAS questionnaire yielded two factors that accounted for nearly 50% of the variance: foreign language anxiety and learning anxiety. Learning anxiety has been reported to hinder learning outcomes in that individuals who possess learning anxiety may "attend to fewer environmental cues, encode information less well, process material less effectively, experience more cognitive interference, and lose working memory capacity by worrying" (Mueller, 1992; Warr & Downing, 2000, p. 317). Such a description of this factor score aligns well with the items that loaded onto the learning anxiety factor score: "I often feel like not going to my language class" and "During language class, I find myself thinking about things that have nothing to do with the course." This two-factor result is somewhat in line with previous factor analyses, namely those of Horwitz et al. (1986), who identified communication apprehension, social-evaluative anxiety, and test anxiety as the three underlying factors present within foreign language classroom anxiety. MacIntyre and Gardner (1989) uncovered two factors: general anxiety and communicative anxiety in 1989; in 1991, MacIntyre and Gardner extracted three: social evaluation, communication apprehension, and test anxiety.

Two factors emerged from the MASQ and accounted for 37% of the variance – positive emotionality and somatic distress (Reidy & Keogh, 1997; Seidenberg et al., 1994). Previous studies, however, have identified a three-factor model of positive affect, anxious arousal, and general distress (Keogh & Reidy, 2000). Consistent with previous research (Fresco et al., 2002), the PSWQ resulted in one factor, worry, which accounted for 51% of the variance. Also in line with previous research, two factors emerged from the STICSA analysis, cognitive distress and somatic distress, representing nearly 41% of the variance (Gros et al., 2007).

To confirm the reliability of the initial exploratory factor analyses, I performed a secondary confirmatory factor analysis with only the items that loaded strongly onto the factors (cut-off value = .5, [Field, 2013; Pett et al, 2003]). The resulting Cronbach's alpha values were over .7, deemed acceptable for psychological constructs by (Field, 2013). All initial exploratory factor analyses and the subsequent single-factor analysis statistics (i.e., correlation tables and rotated factor loadings) can be found in the appendix of Fox (2016). The mean scores of each anxiety factor were used as predictor variables in ensuing multiple linear regressions to examine their relationship, if any, with vocabulary posttest performance.

Confirmatory factor analysis of self-report variables. The rationale for this procedure was to reduce the number of predictor variables in my regression models in order to isolate the effects of physiology measures without competition from other arguably stronger predictors. Thus, I performed a confirmatory factor analysis (CFA) of a two-factor solution yielding one factor containing all seven anxiety factor scores and another containing all three Delta SAM rating variables. For an in-depth investigation on the impact of the individual anxiety factor scores and mood induction, see Fox et al. (in preparation).

The two-factor model satisfied recommended indices for a good model fit (Hooper, Coughlan, & Mullen, 2008; Schreiber, Nora, Stage, Barlow, & King, 2006).[4] Table 3 displays the fit statistics for the model, and Table 4 provides the loadings of the individual variables within the two-factor model.

Table 3

Goodness-of-fit indicators for the two-factor confirmatory model

X^2	*df*	X^2/df	*p*	*Comparative Fit Index (CFI)*	*Tucker-Lewis Index (TLI)*	*Root Mean Square Error of Approx. (RMSEA)*	*90% Confidence Intervals of RMSEA*
38.44	32	1.20	.20	0.97	0.95	0.05	(0.00, 0.11)

[4] Indices of a good model fit are the following: the χ^2/degrees of freedom ratio be low, the *p*-value of χ^2 >.05, the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) values > 0.95, and the Root Mean Square Error of Approximation (RMSEA) < .07, and its 90% C.I. close to 0 (Hooper et al., 2008; Schreiber et al., 2006).

Table 4

Unstandardized and standardized loadings for the two-factor confirmatory model

Factor	*Item*	*Unstandardized (S.E.)*	*Standardized*	*p-value*
Self-report Anxiety				
	PSWQ Factor 1 – Worry	0.65 (0.13)	5.22	<.001
	STICSA Factor 1 – Cognitive distress	0.68 (0.12)	5.46	<.001
	FLCAS Factor 1 – Foreign language anxiety	0.68 (0.12)	5.50	<.001
	FLCAS Factor 2 – Learning anxiety	0.40 (0.14)	2.94	.003
	MASQ Factor 1 – Positive emotionality	-0.44 (0.13)	-3.38	.001
	MASQ Factor 2 – Somatic distress	0.56 (0.13)	4.44	<.001
	STICSA Factor 2 – Somatic distress	0.48 (0.13)	3.76	<.001
Self-report Delta SAM				
	Valence	0.76 (0.18)	4.18	<.001
	Arousal	-0.53 (0.16)	-3.42	.001
	Dominance	0.28 (0.14)	1.94	.05

Note: Model covariances: Delta Arousal and Delta Valence (estimate = 0.31, S.E. = 0.11, z = 2.81, p = .005); FLCAS Factor 1 and FLCAS Factor 2 (estimate = 0.45, S.E. = 0.12, z = 3.76, p < .001); Self-report SAM and Self-report Anxiety 2 (estimate = -0.50, S.E. = 0.15, z = -3.23, p = .001).

From the two-factor confirmatory model, I extracted the two factor scores and employed them in the ensuing linear multiple regression analyses (Rosseel, 2012).

Vocabulary scores. The initial posttest scores were selected because I was interested in how anxiety levels impacted vocabulary learning and not longer-term retention of new vocabulary items.

For a comprehensive description of each of the variables, definitions, units of measurement, variable type, and operationalizations, see Appendix J. Table 5 demonstrates the statistical package that was employed for each analysis.

Table 5

Statistical packages employed for each analysis

Analysis	*Package*
Exploratory factor analyses	SPSS, version 23 (Corp, 2014)
Confirmatory factor analyses	*R* (Team, 2016)
RQ 1 Repeated-measures ANOVA	SPSS
RQ 2 Linear multiple regression	SPSS
RQ 3 Hierarchical linear multiple regression	SPSS
RQ 3 Zero-inflated negative binomial regression (receptive posttest)	*R*
RQ 3 Robust standard errors for regressions violating assumption of homoscedasticity	*R*

Research question 1. What physiology measures show clearest changes under the influence of the film-based mood induction?

Self-reports. Three separate 3x2 Repeated Measures ANOVAs were performed to confirm that the emotional manipulation had the expected outcomes on the three dimensions of emotion: valence, arousal, and dominance. The within-subjects variable was time (3 levels: Time 1 [start of day 1 experiment], 2 [immediately following the film mood induction], and 3 [end of day 1 experiment]), and the between-subjects variable was mood induction (2 levels, neutral and negative). Results of these analyses will reveal if there were changes in participants' emotional self-report as a function of time, mood induction group, and/or a combination of time and mood induction group.

Physiology. I performed three separate 6x2 RM ANOVAs for each dependent variable: HR, HRV, and SCL. The within-subject independent variable was time (6 levels, one for each of the baseline-corrected active films), and the between-subjects independent variable was mood induction (2 levels, neutral and negative). Results of these analyses uncovered significant main effects of time and mood induction. At the end of each physiological variable, I will explain how I carried forward the specific physiological measure into the ensuing regression analyses.

Research question 2. How does a film-based mood induction influence vocabulary learning on a paired-associates learning task? At the outset of the analyses, I anticipated employing the forced-entry approach with the reverse-selection model-trimming method (Crawley, 2007; Larson-Hall, 2009); however, based upon preliminary analyses of the data, this approach did not yield an informative results model. One possible reason for this was that too many predictor variables were entered into the initial regression model. Howell (2012) recommended a bare minimum of ten participants per predictor variable. Clearly, entering in the variables for mood induction, the seven anxiety factor scores, interaction terms between mood induction and the seven anxiety factor scores, the two Delta SAM self-reports, the three physiology measures, and their interactions with mood induction (total = 23), flooded the model. Despite the intention of and attempts to perform the reverse-selection method in order to arrive at the most parsimonious model, the physiology measures revealed no predictive influence. For that reason, I performed hierarchical regression models to identify individual variables' contributions to the overall models.

In order to preserve power of my statistical model, I performed a single variable linear multiple regression to observe the influence of treatment upon the vocabulary posttests. Furthermore, I performed correlation analyses of the outcome and predictor variables to observe initial patterns in the data.

Research question 3. What are the effects of physiology on vocabulary learning on a paired-associates learning task (a) in isolation, and (b) when controlling for the mood induction, state-level self-report measures, and self-report measures of trait-level anxiety? To answer part A of RQ 3, I will enter three physiology variables extracted from the main effects of the manipulation checks into a hierarchical linear multiple regression (Larson-Hall, 2009). Due to the lack of relevant guiding research, I determined the most intuitive approach to entering the predictor variables into the hierarchical regression models as such: Baseline HRV, Mean HR, and Delta SCL. My rationale for this variable ordering is based upon previous researchers' findings on each of the measure's propensity to serve as an informative influence of outcome variables. By and large, the variable that may have the least amount of influence is the Baseline HRV, which has predominantly been used as a trait-level indicator for emotion regulation (Park et al., 2013; Thayer et al., 1996). Following Baseline HRV, I entered the Mean HR variable. This measure has been more frequently used compared to Baseline HRV, but there remains a fundamental question regarding this measure's changes and which particular underlying autonomic response the changes represent. As a reminder, both the sympathetic and parasympathetic nerves innervate, or influence, changes in heart rate, which makes it difficult to draw firm conclusions about its behavior (Berntson et al., 2007; Carvalho et al., 2012; Frazier et al., 2004). Finally, I entered the Delta SCL variable last because I expected this to have the most informative power compared to the Baseline HRV and Mean HR. This measure has consistently been employed in psychophysiological studies as a reliable operationalization of arousal (Barry & Sokolov, 1993; Croft et al., 2004; Tremayne & Barry, 2001).

As for Part B of RQ 3, I first entered the mood induction predictor variable in order to determine if mood induction alone had an impact on the participants' outcome. Next, I entered the self-report Delta SAM factor score in order to control for the individual differences associated with state-level changes associated with the participants' experimental group. The third variable included in the model was the self-report anxiety level factor score, which represented a composite of trait-level anxiety. For a more fine-grained report of the influence of anxiety and mood induction on vocabulary learning, see Fox et al. (in preparation). Once I entered the control for mood induction and self-report individual difference, I entered the Delta HRV score, followed by the Mean heart rate, and Delta Skin Conductance, respectively.

In summary, I performed two hierarchical linear multiple regressions on each of the three vocabulary posttests. The first hierarchical regression model isolated the physiological measures for a fine-grained inspection of their influence on vocabulary learning. Next, I controlled for all known effects on vocabulary learning–mood induction, self-reported Delta SAM scores, and self-reported trait-level anxiety scores–and then I introduced the physiology measures to observe whether each physiological measure had an additive and unique effect, and whether or not that effect was significant.

Hypotheses

Research question 1. What physiology measures show clearest changes under the influence of the film-based mood induction?

Self-reports. I hypothesize that the emotional manipulation check for the self-report portion of the study will reveal significant differences between the negative and neutral mood induction group; that is, that the negative group will report lower levels of valence (happiness) and higher levels of arousal (calmness) following the film clips. Furthermore, I posit that both groups will return to comparable levels of valence and arousal by the end of the experiment on day 1 (Fox, 2016; Fox & Miller, 2015; Miller et al., 2017).

Physiology. Based upon previous research regarding physiological responses to emotional stimuli, I expect to observe HR deceleration (Carvalho et al., 2012; Codispoti et al., 2008), decreases in HRV (Demaree et al., 2004; Demaree et al., 2006; Frazier et al., 2004; Kreibig et al., 2007; Pauls & Stemmler, 2003), and increases in SCL for those in the negative mood induction group (Cacioppo & Tassinary, 1990; Carvalho et al., 2012; Codispoti et al., 2008; Kreibig et al., 2007; Palomba et al., 2000). I anticipate the neutral mood induction group to exhibit either no change whatsoever and/or a combination of the following changes: HR deceleration, and decreases in HRV and SCL (Kreibig et al., 2007).

Research question 2. How does a film-based mood induction influence vocabulary learning on a paired-associates learning task? I expect that negative mood induction will have an overall detrimental effect on all vocabulary posttests (Fox, 2016; Fox & Miller, 2015; Miller et al., 2017).

Research question 3. What are the effects of physiology on vocabulary learning on a paired-associates learning task (a) in isolation, and (b) when controlling for the mood induction, state-level self-report measures, and self-report measures of trait-level anxiety? For the physiology variables in isolation, I expect that participants with high trait levels of HRV and SCL to perform better (Dawson et al., 2007; Frazier et al., 2004; Huang-Pollock et al., 2002; Jönsson, 2007; Thayer et al., 2009; Thayer & Lane, 2009). For the full model regression, I anticipate that a blend of mood induction, self-report, and physiology variables will significantly account for variability in the vocabulary posttests. This hypothesis is based upon the findings of previous research and preliminary analyses (Fox, 2016; Fox et al., in preparation; Miller, 2016; Miller et al., 2017).

CHAPTER 4

Results

Research question 1: Manipulation checks. What physiology measures show clearest changes under the influence of the film-based mood induction?

Six analyses of variances (ANOVAs) were performed to investigate changes in self-reported valence, arousal, and dominance levels as indicated on the self-assessment manikin (SAM) and changes in physiological responses during the emotional induction.

Self-report.

Valence. Table 6 displays the means and standard deviations of the two groups' reported valence and arousal over the three time periods. A 3 (Time) x 2 (Mood induction group) Repeated Measures ANOVA was performed on the participants' reported valence and arousal measures.[5] The within-subjects variable was time (3 levels: Time 1 [start of the experiment], Time 2 [immediately following the film mood induction], and Time 3 [end of day 1 experiment]), and the between-subjects variable was mood induction (2 levels: neutral and negative mood induction).

[5] There is debate surrounding the mood induction of Likert-scale data as continuous variables. According the research that I accessed, Likert data may be used in a parametric procedure (i.e., as continuous variables) when there are five or more categories to the Likert scale (Grace-Martin, 2008). Moreover, the appropriateness of their use is corroborated by the strong significance of the findings ($p < .001$). Had the significance level been closer to the .05 or .01 threshold (for a stricter standard), then treating Likert-scale items as continuous variables may be considered to be bending statistical assumptions too greatly.

Table 6

Emotional self-report manipulation check

Emotion	Mood induction group (*n*)	Time 1 (initial)		Time 2 (post-Tx)		Time 3 (end day 1)	
		M	SD	M	SD	M	SD
Valence	Neutral (35)	7.17	1.22	7.29	1.36	6.17	1.74
	Negative (35)	7.29	1.13	3.80	1.59	6.09	1.46
Arousal	Neutral (35)	3.60	1.59	4.09	1.77	4.29	2.12
	Negative (35)	3.46	1.93	4.97	2.42	3.86	1.82
Dominance	Neutral (35)	5.60	1.61	5.43	1.79	5.51	2.02
	Negative (35)	5.14	1.40	4.66	1.61	5.34	1.47

Note: Emotion self-report was on a Likert scale from 1 to 9.

For valence, the participants' level of happiness, there was a significant time*mood induction interaction (Huynh-Feldt adjustment, $F_{1.91,\ 130.13} = 52.51$, $p < .001$, partial eta-squared = .44, power = 1.0). The Huynh-Feldt adjustment was utilized because Mauchly's test indicated that the assumption of sphericity was violated, $\chi(2) = .911$, $p = .045$. In order to determine when the significant differences occurred, I ran post-hoc tests between the groups' mean valence scores at each time period. I employed the Bonferroni adjustment for multiple comparisons. Significant group means differences were only present at Time 2 (mean difference = 3.49, $p < .001$, 95% CI [2.78, 4.19]).

Further pairwise comparisons revealed significant changes within groups at all three times except for the neutral group from Time 1 to 2 (mean difference = 0.11, $p = 1.0$, 95% CI [-.60, .83]). This means that only the negative group reported significantly different levels of happiness from the time they began the experiment to immediately following the film exposure, and from the time immediately following the film exposure to the end of the experiment. Table 7 displays the mean difference scores for the significant pairwise comparisons, and *Figure 9* displays the interaction plot.

Table 7

Pairwise comparisons of mood induction by time interaction of mean valence levels

Time	*Mood induction group*		*Mean difference*	*Significance (p-value)*	*95% Confidence Interval*
1	Neutral	Negative	-0.11	.69	-0.68, 0.45
2	Neutral	Negative	3.49	<.001	2.78, 4.19
3	Neutral	Negative	0.09	.82	-0.68, 0.85

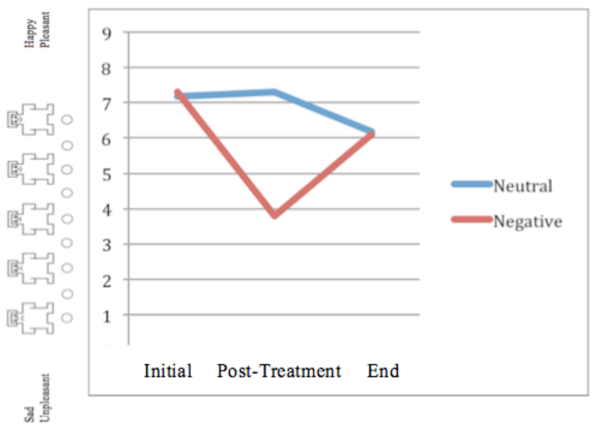

Figure 9. Interaction effect of time and mood induction on change in valence scores by group

Arousal. As for arousal, the participants' level of excitement, there was a significant time*mood induction interaction ($F_{2,\ 136}$ = 3.67, *p* = .03, partial eta-squared = .05, power = .67). No estimated Epsilon adjustments were required in this analysis because Mauchley's test of sphericity was satisfied (*W* = .954, *p* = .203). Pairwise comparisons revealed that the statistical differences were located within the negative mood induction group between Time 1 and 2 (mean difference = 1.51, *p* = .001, 95% confidence interval [2.48, .55]) and between Time 2 and 3 (mean difference = -1.11, *p* = .009, 95% confidence interval [-2.0, -.23]). This indicates that only the negative group displayed significant changes in arousal level immediately following the exposure to the film clips. Furthermore, the significant change in arousal level between Time 2 and 3 suggests that those in the negative mood induction group were able to recover, or return to their initial arousal levels, by the end of the experiment. *Figure 10* displays the interaction effect of time and mood induction on the change of arousal scores by group.

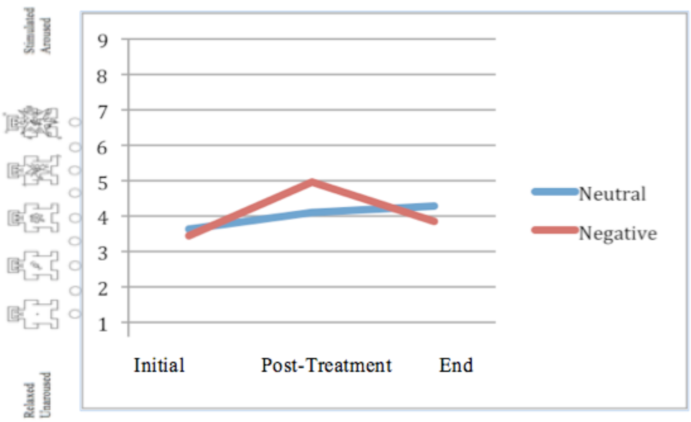

Figure 10. Interaction effect of time and mood induction on change in arousal scores by group

Dominance. As for dominance, there was a nearly significant main effect of time (Huynh-Feldt adjustment, $F_{1.92, 130.20} = 3.08$, $p = .052$, partial eta-squared = .04, power = .57), but no significant effects of mood induction (Huynh-Feldt adjustment, $F_{1, 68} = 1.81$, $p = .18$, partial eta-squared = .03, power = .26) or time*mood induction (Huynh-Feldt adjustment, $F_{1.92, 130.20} = 1.60$, $p = .21$, partial eta-squared = .02, power = .33). The Huynh-Feldt adjustment was utilized because Mauchly's test indicated that the assumption of sphericity was violated, $\chi(2) = .91$, $p = .046$. *Figure 11* displays the interaction effect of time and mood induction on the change of dominance scores by group.

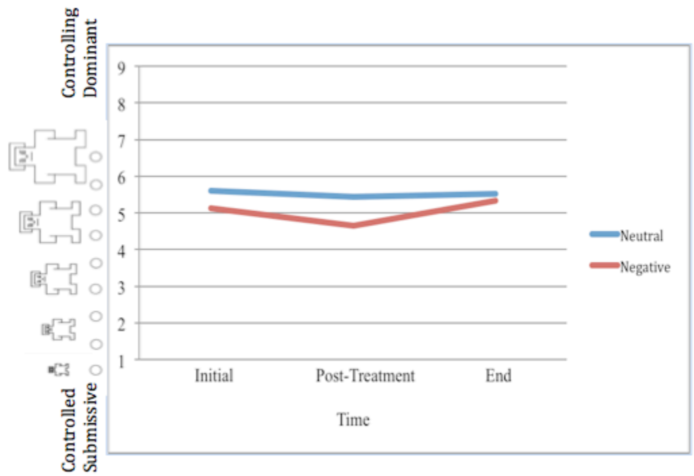

Figure 11. Interaction effect of time and mood induction on change in dominance scores by group

Psychophysiology. As described in the methodology chapter, each physiological measure consisted of six baseline-corrected mean scores representing the six 40-sec films of the active phase of the emotional mood induction. Each physiological measure was entered into a 6 (Time) x 2 (Mood induction group) Repeated Measures ANOVA. Results are reported for each measure below.

Heart rate. The total number of cases analyzed was 47 (neutral n = 27, negative n = 20); initially, 17 participants' heart rate data were unable to be read by the Brain Vision Analyzer (Alhabash, 2016; Analyzer, 2006; Moser, 2016). In order to satisfy the assumption of normality of the dependent variable, I performed two rounds of univariate and multivariate outlier identification and elimination. The systematic approach I employed to identify and eliminate univariate outliers was as follows: I transformed the raw HR means scores into z scores and then eliminated any participant whose data fell outside the recommended range of +/- 3.29 SD (Williams, 2016). This is the same protocol I employed for the following measures, HRV and SCL. As for the multivariate outliers, I calculated the Mahalanobis score for each of the observations and eliminated any that exceeded the critical χ^2 value at the α significance level of

.001 (Heck, 2016). Based upon these diagnostics, I eliminated six participants' data.[6] The subsequent tests of normality indicated that the assumption of normality was satisfied ($p > .05$) except for the negative group at the fourth film (Kolmogorov-Smirnov = .21, $p = .02$; Shapiro-Wilk = .89, $p = .02$).

Table 8 displays the descriptive statistics of the mean baseline-corrected heart rate change scores. As a reminder, the heart rate change scores are the changes in heart rate from the preceding resting phase to the active phase.

Table 8

Descriptive statistics for mean change scores in heart rate by group

Time / Film number	*Mood induction Group (n)*	*Mean Difference*	*SD*
1	Neutral (27)	-1.23	4.54
	Negative (20)	-5.36	6.85
2	Neutral (27)	-1.23	3.60
	Negative (20)	-1.38	3.25
3	Neutral (27)	-0.10	4.20
	Negative (20)	-3.27	5.03
4	Neutral (27)	0.71	5.21
	Negative (20)	-2.03	4.64
5	Neutral (27)	-1.10	4.58
	Negative (20)	-3.77	4.42
6	Neutral (27)	-1.26	4.33
	Negative (20)	-0.75	4.51

Results of the Repeated Measures ANOVA revealed a significant main effect of mood induction, a nearly significant main effect of time, and no time * mood induction interaction. See Table 9 for a summary of the results.

[6] I eliminated case numbers 211, 219, 220, 303, 314, and 315.

Table 9

Repeated-measures ANOVA results for heart rate mean difference scores

	$F_{(4.70, 211.30)}$	*p*	η_p^2	*Power*
Time	2.16	.06	.05	.68
Time x Mood induction	1.87	.11	.04	.61
Mood induction	10.84	.002	.19	.90

Note: p <.05, Huynh-Feldt adjustment reported because Mauchly's Test of Sphericity was significant (p = .027). Huynh-Feldt was chosen because the value (.94) exceeded of the Epsilon estimation the 0.75 cut-off.

Figure 12 displays the profile plot for the baseline-corrected heart rate scores by group in the repeated-measures ANOVA. The baseline-corrected heart rate mean scores of those in the negative mood induction group were consistently lower than those in the neutral group.

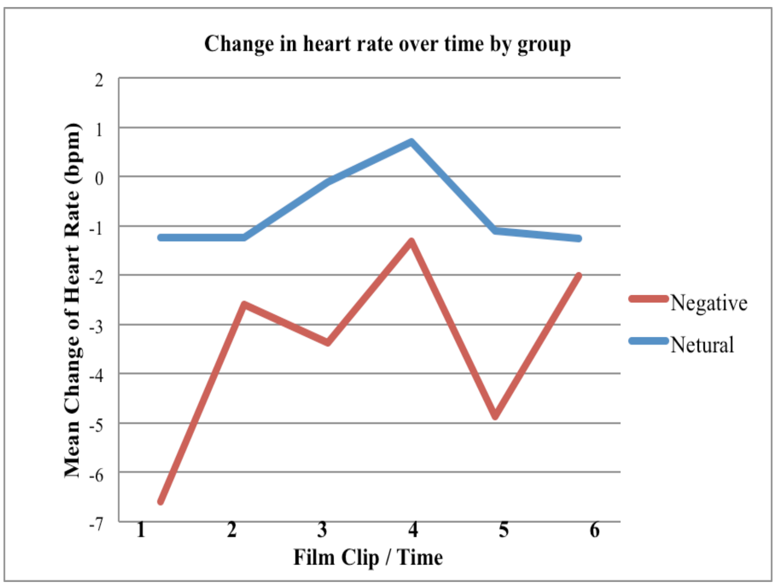

Figure 12: Heart rate change scores over time by group

For the ensuing analyses, I captured the significant effect of mood induction by mean aggregating the six mean change scores into one mean difference score for heart rate.

Heart rate variability. Similarly to the HR analyses, I inspected the raw baseline-corrected HRV scores using descriptive statistics. Initial tests of normality and visual inspection of the histograms and residuals pointed to the presence of influential outliers. Similar to the HR data, a number of cases were missing due to technical issues with the Brain Vision Analyzer software. In addition to missing data, initial inspection of the normality of this variable indicated the HRV data were not normally distributed. I performed the same univariate and multivariate diagnostics as explained for the HR analysis, and eliminated 9 additional participants' data, resulting in a total of 44 participants (neutral $n = 25$, negative $n = 19$).[7] Table 10 displays the descriptive statistics for mean changes in heart rate variability by group.

Table 10

Descriptive statistics for mean changes in heart rate variability by group

Time / Film number	*Mood induction Group (n)*	*Mean Difference*	*SD*
1	Neutral (25)	-74.22	1416.42
	Negative (19)	-2166.14	5959.99
2	Neutral (25)	171.89	786.67
	Negative (19)	499.84	2068.72
3	Neutral (25)	317.90	1614.55
	Negative (19)	-532.14	5611.83
4	Neutral (25)	848.43	1843.10
	Negative (19)	1210.39	3121.93
5	Neutral (25)	645.25	1753.23
	Negative (19)	8.72	1694.23
6	Neutral (25)	34.85	724.17
	Negative (19)	-836.14	5861.84

Note: The mean differences have a large distribution, and this is due to the fact that these are raw difference scores with no transformation applied to that data.

Results of the RM ANOVA revealed a significant effect of time but no main effect of mood induction or time * mood induction interaction. Table 11 displays the results for the Repeated-measures ANOVA for Heart Rate Variability Mean Difference Scores.

[7] I eliminated case numbers 201, 212, 228, 229, 230, 302, 312, 314, and 329. One of these eliminated cases overlapped with the eliminated cases for the previous HR analyses, and two of the eliminated cases overlapped with the upcoming SCL analyses.

Table 11

Repeated-measures ANOVA results for heart rate variability mean difference scores

	$F_{(4.02,\ 169.00)}$	p	η_p^2	*Power*
Time	2.44	.05	.06	.69
Time x Mood induction	0.94	.44	.02	.29
Group	2.58	.12	.06	.35

Note: $p < .05$, Huynh-Feldt adjustment reported because Mauchly's Test of Sphericity was significant ($p < .001$). Huynh-Feldt was chosen because the value (.81) exceeded of the Epsilon estimation the 0.75 cut-off.

Figure 13 offers a visual representation of the change in mean heart rate variability over time by group.

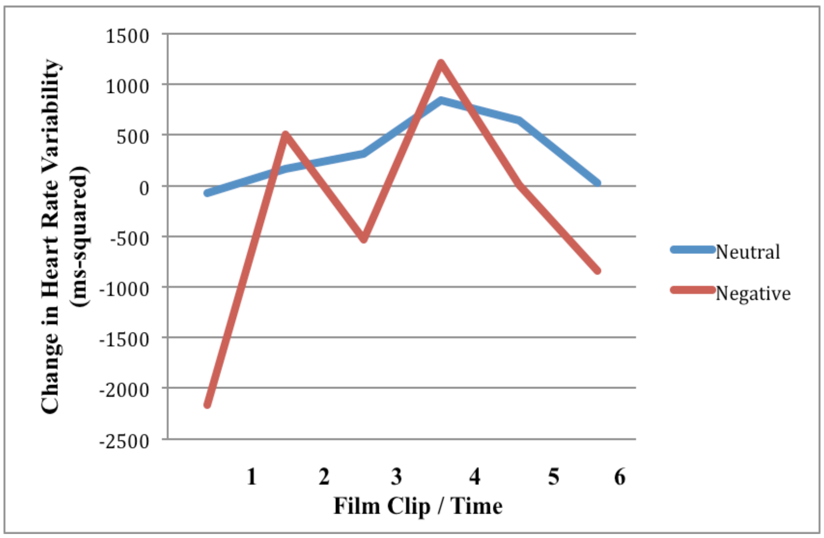

Figure 13. Change in heart rate variability scores over time by group

In order to determine where the significance in time resided, I performed post-hoc pairwise comparisons; however, no significant differences between time segments across both groups were found.[8] Despite the fact that there was a borderline significant effect of time, I chose to carry the Baseline HRV measure forward into the ensuing regression analyses. As a reminder, the Baseline measure is the 90-second time during the two practice film clips and 10-second blank screen. My rationale for selecting this measure over the borderline significant main effect

[8] I suspect this lack of information is due to the fact that I used the Huynh-Feldt adjustment for the F-value.

of time was twofold. First, the pairwise analyses did not reveal where the significance of time was located, and this is most likely due to the non-linear nature of the curve shown in *Figure 13*. Secondly, previous research has primarily focused upon the impacts of trait-level HRV measures as an indicator of an individual's capacity for emotional regulation; generally, those with higher levels of high frequency HRV tend to perform better on tasks under pressure (Park et al., 2014; Thayer et al., 2010).

Skin conductance levels. The skin conductance level was analyzed using the same time segments as the HR and HRV measures. At the outset, three participants' data were missing. After univariate and multivariate diagnostic testing for outliers, I eliminated 10 participants' data, resulting in a total number of 57 (neutral n = 32, negative n = 25).[9] Table 12 displays descriptive of the baseline-corrected SCL scores by group.[10]

Table 12

Descriptive statistics for mean delta skin conductance levels by group

Time / Film number	*Mood induction Group (n)*	*Mean Difference*	*SD*
1	Neutral (32)	1161.71	1758.49
	Negative (25)	1104.03	894.81
2	Neutral (32)	-74.45	648.30
	Negative (25)	115.59	836.01
3	Neutral (32)	-209.77	973.89
	Negative (25)	93.49	908.32
4	Neutral (32)	120.42	917.98
	Negative (25)	340.59	712.94
5	Neutral (32)	446.09	1403.42
	Negative (25)	-25.36	1083.31
6	Neutral (32)	-3.98	643.15
	Negative (25)	485.33	1174.80

[9] I eliminated case numbers 209, 210, 211, 212, 220, 221, 222, 319, 328, and 302. Three of these cases overlapped with the eliminated cases from the HR and HRV analyses.

[10] One may notice how wide the range of scores for this measure is. Although previous research has consistently transformed data to satisfy the assumption of normality, I determined that using the raw scores resulted in a better distribution of the dependent variables and their residuals compared to the transformed ones.

Similar to the HR and HRV measures, a 6 (Time) x 2 (Mood induction group) Repeated Measures ANOVA was performed, which yielded a significant main effect of time ($F_{(4.48, 246.13)}$ = 9.22, *p* = < .001, partial eta-squared = .14, power = 1.00). The main effect of mood induction and the time * mood induction interaction were not significant. Results of the RM ANOVA are reported in Table 13.

Table 13

Repeated-measures ANOVA results for skin conductance level mean difference scores

	F	*p*	η_p^2	*Power*
Time	9.22	<.001	.14	1.00
Time x Group	4.48	.22	.03	.47
Group	0.96	.33	.02	.16

Note: *p < .05, these values are based upon the Huynh-Feldt adjustment due to the violation of the assumption of sphericity (χ = 31.02, *p* = .006). Huynh-Feldt was chosen because the value (.90) exceeded of the Epsilon estimation the 0.75 cut-off.

Post-hoc pairwise comparisons confirmed that the significant differences across both groups occurred between Time 1 (i.e., the first film clip) and the remaining film clips. Table 14 lists the pairwise comparisons of the main effect of time for the skin conductance level Repeated-measures ANOVA.

Table 14

Pairwise comparisons of main effect of time of skin conductance level changes

Time difference	*Mean Difference*	*Std. Error*	*p-value*	*95% Confidence Interval*
1 to 2	3.81	1.12	.02	0.39, 7.23
1 to 3	4.60	1.04	.001	1.41, 7.80
1 to 4	2.86	0.95	.06	-0.70, 5.78
1 to 5	3.62	0.95	.01	0.70, 6.53
1 to 6	3.51	1.07	.03	0.23, 6.78

Note: Based on estimated marginal means; adjustment for multiple comparisons: Bonferroni.
Figure 14 displays the change in skin conductance levels over time by group.

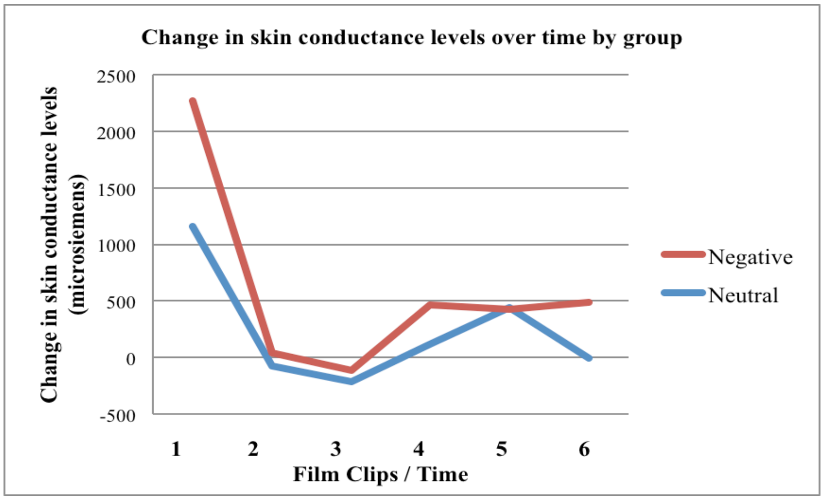

Figure 14. Mean delta skin conductance over time by group

In order to carry this significant main effect of time into the second research question, I created a change score of SCL mean differences by subtracting Time 6 from Time 1. The rationale for this decision was to capture the change over time that best represented the shape of the habituation curve.

In summary, the results of the physiological manipulation checks revealed that there were significant changes by time (i.e., HRV and SCL) and mood induction group (i.e., HR). Table 15 is a summary of the significant effects of the manipulation checks and how I plan to carry each variable into the ensuing regression analyses.

Table 15

Significant effects from the physiology manipulation checks

Measure	*Main effect*	*p*	*Variable carried into ensuing analyses*
Heart Rate (HR)	Mood induction	.002	Mean HR = mean score of HR x 6 Time
Heart Rate Variability (HRV)	Time	.05	Baseline HRV
Skin Conductance Levels (SCL)	Time	<.001	Delta SCL = Time 1 – Time 6

Research questions 2 and 3: The effect of the mood induction and psychophysiology on vocabulary learning on a paired-associates learning task.

A preliminary step to the multiple regression analyses for RQs 2 and 3 was checking the assumptions for linear multiple regression which includes normal distribution of data (including outliers) and equality of variances (Larson-Hall, 2009). I ran the full model for the free recall and productive posttests; diagnostics were not performed on the receptive posttest variable due to the violation of the assumption of normality, which eventually led to my running a zero-inflated negative binomial regression on the variable.

Table 16 offers a summary of the critical statistics for checking assumptions of normality.[11] The values reported in the table represent the second round of assumption diagnostics I performed following the elimination of two influential outliers that yielded high Mahalanobis distance values across both posttests. Those participants were excluded from all further analyses, which reduced the total number of participants to 68. However, due to missing data from physiology measures, 22 participants' data in total was excluded, meaning a total of 48 participants' data were used in the subsequent regression analyses.

Table 16

Summary statistics of assumption diagnostics

	Std. Residual		*Mahalanobis distance*		*Cook's distance*		*Kolmogorov-Smirnov (sig.)*	*Shapiro-Wilks (sig.)*
	Mean (SD)	*Min., Max.*	*Mean (SD)*	*Min., Max.*	*Mean (SD)*	*Min., Max.*		
Free Recall	0.00 (0.97)	-1.85, 2.64	2.94 (3.89)	0.04, 20.73	.02 (.03)	<.001, .18	.06 (.20)	.98 (.68)
Productive	0.00 (0.94)	-1.99, 1.83	5.88 (4.19)	1.38, 21.37	.02 (.03)	<.001, .10	.10 (.20)	.98 (.60)

---

[11] Assumptions of normality are satisfied when the following conditions are met: standard residuals do not exceed or fall short of 3 standard deviations from the mean, Cook's distance values do not exceed 1.00, and Mahalanobis distance values exceeding 15 for large sample sizes (>30) are further investigated (Larson-Hall, 2009).

The final assumption for multiple linear regressions is homogeneity of variances. I examined the scatterplots between the studentized residuals and predicted values of the standardized residuals for each step of the physiology-only and full regression models. Multiple models did not display the expected cloud of randomly distributed data, so I performed robust regressions on the models that violated this assumption and reported the accompanying *B* estimate, adjusted standard errors, and *p*-values (Larson-Hall, 2009).

Descriptive statistics. Table 17 displays the descriptive statistics for the three vocabulary posttests. A general trend that can be observed in the vocabulary posttests is that the mean scores of the tests incrementally increased from the Free Recall, to Productive, and Receptive posttests; furthermore, the neutral group outperformed the negative group on each of the posttests. One possible reason why this trend emerged was due to increased exposure to the target words during the course of the preceding posttests; others would argue that the tests were presented from most to least difficult.

Table 17

Descriptive statistics of vocabulary posttests

Vocabulary Test type	*Group*	*Mean (SD)*	*Min.*	*Max. (out of 24)*	*CI*	*Kolmogorov-Smirnov Test of Normality (p)*
Free Recall	Neutral (*n* = 33)	10.81 (2.49)	6	16.63	9.93, 11.70	.07 (.20)
	Negative (*n* = 35)	9.68 (3.58)	3.5	17	8.44, 10.90	.08 (.20)
Productive	Neutral (*n* = 33)	16.21 (4.87)	5.25	23.63	14.48, 17.94	.14 (.10)
	Negative (*n* = 35)	12.95 (6.76)	2.25	24	10.63, 15.27	.11 (.20)
Receptive	Neutral (*n* = 33)	19.83 (4.14)	10	24	18.36, 21.30	.20 (.001)
	Negative (*n* = 35)	16.37 (6.34)	4	24	14.19, 18.55	.12 (.20)

Table 18 displays all the predictor variables entered into the multiple regressions.

Table 18

Summary statistics of predictor variables

Predictor Variables	*n*	*Mean*	*Std. Deviation*	*Min.*	*Max.*
Mood induction	68	0.00	1.00	0.51	0.50
Self-report Delta SAM	68	-1.00	2.00	0.00	0.84
Self-report Anxiety	68	-1.00	2.00	0.02	0.89
Baseline Heart Rate Variability (ms^2, baseline corrected)	48	59.00	18126.00	2755.04	4275.81
Mean Heart Rate (bpm, baseline corrected)	50	-13.00	7.00	-1.51	3.26
Delta Skin Conductance Level (μS, baseline corrected)	65	-3381.00	6536.00	833.97	1808.01

Research question 2: How does a film-based mood induction influence vocabulary learning on a paired-associates learning task?

Table 19 displays the results of the mood induction variable in isolation with each vocabulary posttest. Based upon the results of these initial regressions (*n* = 68), the negative mood induction had a negative influence on all vocabulary posttests, and significant negative influence on the productive and receptive (count portion[12]) posttests.

[12] This analysis will be more fully described in the RQ 2 and 3 results section for the receptive posttests. In particular, this portion of the zero-inflated negative binomial regression modeled errors, which explains why the count model coefficient (*B* = 0.47) is positive in this case.

Table 19

Summary of regression analyses of mood induction on all vocabulary posttests

				Unstandardized Coefficients		*Std. Coefficients*		
Posttest	*Variables*	R^2	*Adj.* R^2	*B*	*Std. Error*	*β*	*95% C.I.*	*p*
Free recall		.03	.02					
	(Constant)			10.81	0.54		9.73, 11.89	<.001
	Mood induction			-1.14	0.75	-0.18	-2.64, 0.37	.14
Productive		.072	.058					
	(Constant)			16.21	1.03		14.15, 18.27	< .001
	Mood induction			-3.26	1.44	-0.27	-6.12, 0.39	.03
Receptive								
				Unstandardized Coefficients		*Standardized Coefficients*		
				B	*Std. Error*	*z*		*p*
				Count model coefficients (neg. bin. with log link)				
	(Constant)			2.47	0.14	18.19		< .001
	Mood induction			0.47	0.18	2.63		.008
	Log(theta)			1.09	0.25	4.33		< .001
				Zero-inflation model coefficients (binomial with logit link)				
	(Constant)			-0.86	0.39	-2.23		.03
	Mood induction			-0.54	0.58	-0.94		.35

Note: Constant = neutral mood induction group

To further explore the relationship of mood induction with other predictor variables, I examined the correlation table show in Table 20. Variables that showed strong relationships with the mood induction variable were Self-report Delta SAM (r = .64 , p < .001), Self-report Anxiety (r = -.27, p = .03), and Mean Heart Rate (r = -.48, p < .001). These correlations reveal that the mood induction is also related to measures of physiology and self-reports, and that is an important consideration in the ensuing regressions.

Table 20

Correlation table for three posttests and predictor variables

	Free Recall	Productive	Receptive	Mood induction	Self-report Delta SAM	Self-report Anxiety	BL HRV	Mean HR	Delta SCL
Free Recall (*n* = 68)	1								
Productive (n = 68)	.87 (<.001)	1							
Receptive (n = 68)	.75 (<.001)	.89 (<.001)	1						
Mood induction (n = 68)	-.18 (.14)	-.27 (.03)	-.31 (.01)	1					
Self-report Delta SAM (n = 68)	-.16 (.19)	-.20 (.10)	-.25 (.04)	.64 (<.001)	1				
Self-report Anxiety (n = 68)	-.02 (.87)	-.02 (.87)	.01 (.94)	-.27 (.03)	-.63 (<.001)	1			
Baseline HRV (n = 48)	.08 (.58)	.12 (.41)	.08 (.58)	.26 (.08)	-.001 (.99)	.10 (.51)	1		
Mean HR (n = 50)	.26 (.06)	.25 (.08)	.34 (.02)	-.48 (<.001)	-.24 (.10)	.02 (.88)	-.24 (.10)	1	
Delta SCL (n = 65)	.20 (.10)	.16 (.19)	.19 (.13)	-.05 (.68)	-.21 (.09)	.20 (.11)	-.25 (.09)	.12 (.40)	1

Research question 3: What are the effects of physiology on vocabulary learning in a paired-associates learning task (a) in isolation, and (b) when controlling for mood induction, self-reported measures of state-level emotions, and self-reported measures of trait-level anxiety?

Free recall posttest.

Part A: Physiology only. The first variable entered into the model was Baseline HRV, and it resulted in no explanation of variance. The second model included the Mean HR variable, which proved to be a significant predictor (B = 0.27, β = 0.31, p = .02). In the final model, both the Mean HR (B = 0.25, β = 0.22, p = .01) and Delta SCL (B = 0.00, β = 0.33, p = .004) variables significantly accounted for variation. The size of the unstandardized B reflects the unit of measure, micro-Siemens, which is a very small measure. To gain a better understanding of the relative impact of Delta SCL compared to the other predictor variables, I turn to the standardized Beta (β). In the final model, the standardized Beta is 0.33, suggesting that Delta SCL accounted for a larger portion of the variance than any of the other variables.

The Delta SCL finding indicates that participants who had the larger Delta skin conductance levels from Time 1 to Time 6 – those who may have been initially highly aroused but then recovered well – performed better on the vocabulary posttests than those with relatively little variation in their skin conductance levels over the course of the film-based mood induction. Thus, regardless of whether a participant was in the negative or neutral mood induction group, those with larger delta skin conductance levels performed better on the free recall posttest. As for the significant effect of the Mean HR variable, those with a higher mean level of HR also performed better on the vocabulary posttest. Table 21 displays the summary of the physiology-only regression hierarchical model for the free recall posttest.

Table 21

Summary of physiology-only hierarchical regression analysis for variables predicting free recall vocabulary posttest scores (n = 50)

						Coefficients				
						Unstandardized		Std.		
Model	*Variables*	R^2	*Adj.* R^2	ΔR^2	*p-value of F-test*	*B*	*Std. Error*	β	*95% C.I.*	*p*
1		.007	-.02	.007	.58					
	(Constant)					10.08	0.52		9.44, 11.15	<.001
	Baseline HRV					0.00	0.00	0.08	0.00, 0.00	.73
2		.10	.06	.09	.04					
	(Constant)					10.38	0.51		9.71. 11.52	<.001
	Baseline HRV					0.00	0.00	0.16	0.00, 0.00	.51
	Mean HR					0.27	0.11	0.31	-.02, 0.47	.02
3		.24	.19	.14	.006					
	(Constant)					9.72	0.51		9.20, 11.09	<.001
	Baseline HRV					0.00	0.00	0.04	0.00, 0.00	.24
	Mean HR					0.25	0.09	0.22	-0.05, 0.42	.01
	Delta SCL					0.00	0.00	0.33	0.00, 0.001	.004

Note: Constant = neutral mood induction group; robust standard errors were calculated due to evidence of heteroskedasticity of the residuals.

Part B: Full model. As previously explained in chapter 3, the full model hierarchical regression was constructed following this entry method: mood induction, self-report Delta SAM, self-report Anxiety, Baseline HRV, Mean HR, and Delta SCL. The final model revealed Delta SCL (B = 0.00, β = 0.43, p = .003) significantly contributed to vocabulary learning even when all other variables were controlled. Furthermore, the standardized Beta (β = 0.43) once again indicates that this variable accounted for a large share of the variance. Table 22 displays the summary of the full model regression for the free recall posttest.

Table 22

Summary of full model hierarchical regression analysis for variables predicting free recall vocabulary posttest scores (n = 48)

						Unstandardized Coefficients		*Std. Coefficients*		
Model	*Variables*	R^2	*Adj.* R^2	ΔR^2	*p-value of F-test*	*B*	*Std. Error*	*β*	*95% C.I.*	*p*
1		.07	.05	.07	.07					
	(Constant)					10.95	0.56		9.82, 12.07	< .001
	Mood induction					-1.49	0.81	-0.26	-3.11, 0.14	.07
2		.07	.03	.002	.79					
	(Constant)					10.83	0.71		9.40, 12.26	< .001
	Mood induction					-1.26	1.17	-0.22	-3.61, 1.10	.29
	Self-report Delta SAM					-0.18	0.68	-0.06	-1.55, 1.19	.79
3		.07	.009	.002	.74					
	(Constant)					10.78	0.73		9.40, 12.25	< .001
	Mood induction					-1.18	1.21	-0.21	-3.61, 1.26	.34
	Self-report Delta SAM					-0.37	0.89	-0.11	-1.53, 1.42	.68
	Self-report Anxiety					-0.20	0.62	-0.07	-1.46, 1.05	.74
4		.10	.01	.02	.30					
	(Constant)					10.72	0.73		9.24, 12.19	< .001
	Mood induction					-1.66	1.29	-0.29	-4.27, 0.94	.21
	Self-report Delta SAM					-0.20	0.90	-0.06	-2.02, 1.62	.83
	Self-report Anxiety					-0.24	0.62	-0.08	-1.50, 1.01	.70
	Baseline HRV					0.00	0.00	0.17	0.00, 0.00	.3

Table 22 (cont'd)

5		.13	.03	.04	.19					
	(Constant)					10.66	0.76		9.19, 12.12	< .001
	Mood induction					-0.96	1.23	-0.17	-3.75, 1.83	.44
	Self-report Delta SAM					-0.27	0.91	-0.08	-2.07, 1.54	.77
	Self-report Anxiety					-0.17	0.57	-0.06	-1.42, 1.08	.76
	Baseline HRV					0.00	0.00	0.18	0.00, 0.00	.44
	Mean HR					0.19	0.14	0.22	-0.10, 0.49	.17
6		.29	.19	.16	.004					
	(Constant)					10.31	0.66		8.94, 11.67	< .001
	Mood induction					-1.96	1.32	-0.35	-4.60, 0.68	.15
	Self-report Delta SAM					0.34	0.92	0.10	-1.37, 2.04	.72
	Self-report Anxiety					-0.16	0.53	-0.05	-1.30, 0.98	.77
	Baseline HRV					0.00	0.00	0.32	0.00, 0.00	.14
	Mean HR					0.14	0.12	0.16	-0.13, 0.41	.25
	Delta SCL					0.00	0.00	0.43	0.00, 0.001	.003

Note: Self-report is a *z*-score, centered at mean = 0; mood induction is for the neutral group; robust standard errors were calculated due to evidence of heteroskedasticity of the residuals.

Productive posttest.

Part A: Physiology only. Similar to the results of the free recall posttest, Mean HR significantly accounted for variation when it was included in the second model ($B = 0.53$, $\beta = 0.31$, $p = .02$). Likewise, Mean HR ($B = 0.49$, $\beta = 0.29$, $p = .01$) and Delta SCL ($B = 0.00$, $\beta = 0.31$, $p = .01$) emerged as significant predictor variables for the productive posttest. The directionality of the coefficients remained the same as in the free recall vocabulary posttest: those with larger Delta SCL levels and higher Mean HR levels performed better on the productive posttest. Table 23 displays the summary of the physiology-only model regression for the productive posttest.

Table 23

Summary of physiology-only hierarchical regression analysis for variables predicting productive vocabulary posttest scores (n = 48)

						Unstandardized Coefficients		*Std. Coefficients*		
Model	*Variables*	R^2	*Adj.* R^2	ΔR^2	*p-value of F-test*	*B*	*Std. Error*	*β*	*95% C.I.*	*p*
1		.02	-.006	.02	.41					
	(Constant)					14.29	0.99		12.31, 16.27	< .001
	Baseline HRV					0.00	0.00	0.12	0.00, 0.001	.45
2		.10	.06	.09	.04					
	(Constant)					14.88	0.96		12.89, 16.87	< .001
	Baseline HRV					0.00	0.00	0.20	0.00, 0.001	.24
	Mean HR					0.53	0.23	0.31	0.03, 1.04	.02
3		.19	.14	.09	.03					
	(Constant)					13.84	0.86		11.70, 15.98	.01
	Baseline HRV					0.00	0.00	0.27	0.00, 0.001	.07
	Mean HR					0.49	0.19	0.29	0.005, 0.98	.01
	Delta SCL					0.00	0.00	0.31	0.00, 0.002	.01

Note: Constant = neutral mood induction group; robust standard errors were calculated due to evidence of heteroskedasticity of the residuals.

Part B: Full model. The first model of the productive hierarchical regression included the significant predictor mood induction variable (B = -3.21, β = -0.29, p = .05) in the expected negative direction. The final model of this regression revealed two significant predictor variables: Baseline HRV (B = 0.00, β = 0.35, p = .02) and Delta SCL (B = 0.00, β = 0.35, p = .01). The standardized Betas for both the Baseline HRV (β = 0.29) and Delta SCL (β = 0.31) indicate that these variables accounted for large shares of the variance. Table 24 displays the summary of the full model regression for the productive posttest.

Table 24

Summary of full model hierarchical regression analysis for variables predicting productive vocabulary posttest scores (n = 48)

						Coefficients				
						Unstd.		*Std.*		
Model	*Variables*	R^2	*Adj.* R^2	ΔR^2	*p-value of F-test*	*B*	*Std. Error*	β	*95% C.I.*	*p*
1		.08	.06	.08	.05					
	(Constant)					16.28	1.10		14.06, 18.50	< .001
	Mood induction					-3.21	1.59	-0.29	-6.42, -0.004	.05
2		.08	.04	.001	.86					
	(Constant)					16.13	1.40		13.2, 18.95	< .001
	Mood induction					-2.92	2.31	-0.26	-7.57, 1.73	.21
	Self-report Delta SAM					1.10	1.34	-0.04	-2.94, 2.47	.86
3		.08	.02	.00	1.00					
	(Constant)					16.13	1.44		13.22, 19.04	< .001
	Mood induction					-2.92	2.39	-0.26	-7.73, 1.89	.29
	Self-report Delta SAM					-0.23	1.75	-0.04	-3.77, 3.30	.89
	Self-report Anxiety					0.00	1.23	0	-2.47, 2.47	1.00
4		.12	.04	.04	.16					
	(Constant)					15.96	1.42		13.07, 18.85	< .001
	Mood induction					-4.21	2.38	-0.37	-9.30, 0.88	.08
	Self-report Delta SAM					0.23	1.93	0.04	-3.33, 3.78	.91
	Self-report Anxiety					-0.11	1.13	-0.02	-2.56, 2.34	.93
	Baseline HRV					0.00	0.00	0.22	0.00, 0.001	.19
5		.15	.05	.03	.24					
	(Constant)					15.85	1.46		12.97, 18.73	< .001
	Mood induction					-2.98	2.71	-0.26	-8.45, 2.50	.28
	Self-report Delta SAM					0.10	1.89	0.02	-3.44, 3.65	.96
	Self-report Anxiety					0.02	1.12	0.004	-2.43, 2.47	.98
	Baseline HRV					0.00	0.00	0.24	0.00, 0.001	.16
	Mean HR					0.34	0.28	0.2	-0.23, 0.92	.22
6		.26	.15	.11	.02					
	(Constant)					15.29	1.29		12.52, 18.05	< .001
	Mood induction					-4.61	2.83	-0.41	-9.97, 0.75	.11
	Self-report Delta SAM					1.09	1.85	0.17	-2.37, 4.55	.56
	Self-report Anxiety					0.04	1.10	0.007	-2.28, 2.36	.97
	Baseline HRV					0.00	0.00	0.35	0.00, 0.001	.02
	Mean HR					0.25	0.23	0.15	-0.30, 0.80	.28
	Delta SCL					0.00	0.00	0.35	0.00, 0.002	.01

Note: Self-report is a z-score, centered at mean = 0; mood induction is for the neutral group (= 0); robust standard errors were calculated due to evidence of heteroskedasticity of the residuals.

Receptive posttest.

Due to a ceiling effect in the receptive posttest, scores for the receptive posttest violated the assumption of normality and unbounded outcome variables. In order to proceed appropriately, I applied a number of manipulations to the data and used a zero inflated negative binomial regression with log transformation (Osborne, 2008).[13] This type of analysis divides the outcome variable in two ways: (a) non-perfect scores (called the count / non-zeros, where zero denotes the number of errors), and (b) the binomial portion comparing the perfect scores (zeros) versus the non-perfect scores (non-zeros). The count portion of the model (a) accounts for those who scored less than perfectly (i.e., non-zeros); because we are modeling errors, a positive regression coefficient means the variable was associated with an increase in errors. The zero-inflation portion of the regression (b) modeled whether any parameters significantly predicted whether participants scored perfectly ($x = 0$) or not ($x > 0$); thus, the intercept (constant) represents the odds of perfect to not perfect performance; therefore, positive B values indicate that the odds of a perfect score went up.

Part A: Physiology only. There were no significant predictor variables in either portion of the zero-inflated negative binomial for the physiology measures. The Mean HR variable approached significance in the final binomial portion of the model ($B = 0.00$, $z = 1.21$, $p = .06$). Table 25 displays the summary of the physiology-only zero-inflated negative binomial regression for the receptive posttest.

---

[13] This is the same analysis performed in Fox et al. (in preparation), which investigated the influence of seven anxiety factor scores and mood induction on the vocabulary posttests.

Table 25

Summary of zero-inflated negative binomial regression analysis for physiology-only variables predicting receptive vocabulary posttest scores (n = 48)

		Unstandardized Coefficients		Standardized Coefficients	
Model	*Variables*	*B*	*Std. Error*	*z*	*p*
1		Count model coefficients (neg. bin. with log link)			
	(Constant)	2.36	0.53	4.44	< .001
	Baseline HRV	0.05	0.08	0.63	.53
	Log (theta)	1.25	0.31	4.07	< .001
		Zero-inflation model coefficients (binomial with logit link)			
	(Constant)	-2.18	1.72	-1.27	.21
	Baseline HRV	0.14	0.24	0.57	.57
2		Count model			
	(Constant)	2.43	0.51	4.76	< .001
	Baseline HRV	0.02	0.07	0.26	.79
	Mean HR	-0.05	0.03	-1.69	.09
	Log (theta)	1.36	0.32	4.32	< .001
		Zero-inflation model			
	(Constant)	-2.47	1.87	-1.32	.19
	Baseline HRV	0.21	0.26	0.82	.41
	Mean HR	0.23	0.12	1.84	.07
3		Count model coefficients			
	(Constant)	2.69	0.54	4.97	< .001
	Baseline HRV	-0.01	0.08	-0.16	.87
	Mean HR	-0.05	0.03	-1.27	.16
	Delta SCL	0.00	0.00	-1.14	.26
	Log (theta)	1.43	0.33	4.37	< .001
		Zero-inflation model			
	(Constant)	-3.95	2.25	-1.76	.08
	Baseline HRV	0.38	0.30	1.27	.20
	Mean HR	0.26	0.14	1.89	.06
	Delta SCL	0.00	0.00	1.21	.23

Part B: Full model. All variables were entered into the full model of the receptive posttest. In the count model portion of the first model, mood induction was a significant predictor (B = 0.47, z = 4.33, p = .008). As a reminder, the regression coefficients modeled errors, so a positive coefficient denotes more errors on the vocabulary posttest. This finding indicates that the negative mood induction increased the probability of making errors for participants who scored less than perfectly on the receptive posttest. There were no other significant predictor variables in this model. Table 26 displays the summary of the full model zero-inflated negative binomial regression for the receptive posttest.

Table 26

Summary of zero-inflated negative binomial regression analysis for full model variables predicting receptive vocabulary posttest scores (n = 48)

		Unstandardized Coefficients		Standardized Coefficients	
Model	*Variables*	*B*	*Std. Error*	*z*	*p*
1		Count model coefficients (neg. bin. with log link)			
	(Constant)	2.47	0.14	18.19	<.001
	Mood induction	0.47	0.18	2.63	.008
	Log (theta)	1.09	0.25	4.33	<.001
		Zero-inflation model coefficients (binomial with logit link)			
	(Constant)	-0.86	0.39	-2.23	.03
	Mood induction	-0.54	0.58	-0.94	.35
2		Count model			
	(Constant)	2.49	0.16	15.2	< .001
	Mood induction	0.44	0.25	1.75	.08
	Self-report Delta SAM	0.02	0.14	0.17	.87
	Log (theta)	1.1	0.25	4.33	< .001
		Zero-inflation model			
	(Constant)	-0.88	0.46	-1.9	.06
	Mood induction	-0.5	0.76	-0.66	.51
	Self-report Delta SAM	-0.04	0.46	-0.08	.93
3		Count model coefficients			
	(Constant)	2.52	0.16	15.27	< .001
	Mood induction	0.37	0.26	1.42	.16
	Self-report Delta SAM	0.14	0.19	0.76	.45
	Self-report Anxiety	0.13	0.14	0.92	.36
	Log (theta)	1.12	0.25	4.4	< .001
		Zero-inflation model			
	(Constant)	-0.96	0.48	-2.01	.04
	Mood induction	-0.35	0.78	-0.45	.65
	Self-report Delta SAM	-0.33	0.6	-0.55	.58
	Self-report Anxiety	-0.34	0.44	-0.78	.44
4		Count model			
	(Constant)	2.29	0.56	4.10	< .001
	Mood induction	0.42	0.32	1.32	.19
	Self-report Delta SAM	-0.06	0.24	-0.24	.81
	Self-report Anxiety	-0.02	0.17	-0.11	.92
	Baseline HRV	0.02	0.08	0.30	.77
	Log (theta)	1.38	0.32	4.35	< .001
		Zero-inflation model			
	(Constant)	-2.33	1.86	-1.25	.21
	Mood induction	-1.50	1.24	-1.21	.23
	Self-report Delta SAM	0.49	0.83	0.60	.55
	Self-report Anxiety	-0.07	0.53	-0.14	.89
	Baseline HRV	0.25	0.27	0.94	.35

Table 26 (cont’d)

5		Count model			
	(Constant)	2.34	0.55	4.24	< .001
	Mood induction	0.32	0.32	0.99	.32
	Self-report Delta SAM	-0.05	0.23	-0.23	.82
	Self-report Anxiety	-0.03	0.16	-0.16	.87
	Baseline HRV	0.01	0.08	0.15	.88
	Mean HR	-0.04	0.03	-1.09	.28
	Log (theta)	1.43	0.32	4.45	< .001
		Zero-inflation model			
	(Constant)	-2.44	1.96	-1.25	.21
	Mood induction	-0.59	1.34	-0.44	.66
	Self-report Delta SAM	0.32	0.82	0.39	.70
	Self-report Anxiety	0.00	0.55	-0.01	.99
	Baseline HRV	0.24	0.28	0.87	.38
	Mean HR	0.20	0.14	1.40	.16
6		Count model			
	(Constant)	2.67	0.58	4.62	.00
	Mood induction	0.41	0.32	1.26	.21
	Self-report Delta SAM	-0.06	0.23	-0.28	.78
	Self-report Anxiety	0.03	0.16	0.18	.85
	Baseline HRV	-0.03	0.08	-0.42	.67
	Mean HR	-0.02	0.03	-0.71	.48
	Delta SCL	0.00	0.00	-1.26	.21
	Log (theta)	1.52	0.34	4.46	.00
		Zero-inflation model			
	(Constant)	-4.09	2.39	-1.72	.09
	Mood induction	-1.11	1.54	-0.72	.47
	Self-report Delta SAM	0.73	0.92	0.80	.43
	Self-report Anxiety	0.10	0.57	0.18	.86
	Baseline HRV	0.46	0.34	1.36	.17
	Mean HR	0.21	0.16	1.31	.19
	Delta SCL	0.00	0.00	1.31	.19

Note: Self-report is a z-score, centered at mean = 0; mood induction is for the neutral group (= 0).

Table 27 displays overall findings from the mood induction (RQ 2), physiology (RQ 3a), and full model (RQ 3b) regressions for each of the three vocabulary posttests. In short, it appears that the Mean HR and Delta SCL were more informative of the vocabulary learning in the physiology-only models; however, only the Delta SCL measure remained a significant predictor variable in the full model. Results will now be discussed in chapter 5.

Table 27

Summary of significant predictor variables of the mood induction, physiology, and full model regression analyses of the free recall, productive, and receptive vocabulary posttests

Posttest	*Mood induction*	*Physiology*	*Full model*
Free recall	Yes	Mean HR Delta SCL	Delta SCL
Productive	No	Baseline HRV Delta SCL	Delta SCL
Receptive	Yes (count portion)	Mean HR (binomial portion)	

CHAPTER 5

Discussion

Manipulation checks

What self-reported and physiology measures showed the clearest changes under the influence of the film-based mood induction?

Self-report.

I hypothesized that participants in the negative group would show greater changes in valence and arousal compared to their neutral group counterparts. Findings confirmed that the negative mood induction had the expected impact: participants who viewed the negative film clips reported significantly lower levels of happiness and significantly higher levels in arousal immediately following the films. The effect size of the time by group interaction for the difference in valence by group (partial eta-squared = .44) is an indication of the strength of this finding. Taken together, self-reports of emotional changes indicate that our emotional manipulation had the intended impact on the two groups of participants; that is, there were more drastic changes reported by those in the negative group compared to those in the neutral group. These findings suggest that participants were conscious of how the mood induction affected their valence and arousal levels, but what remains to be determined is whether the emotional stimuli had measurable and differentiated impacts on the participants' subconscious, involuntary physiological responses. That is the question to which I now turn.

Psychophysiological measures.

The analyses involving the psychophysiological measures revealed both expected and surprising results. For the change in heart rate, there was a significant effect of mood induction; those in the negative mood induction group exhibited a smaller change in heart rate across the six films clips compared to those in the neutral group. As a reminder, the heart is dually innervated–or influenced–by nerves that activate both sympathetic (fight or flight) and parasympathetic (rest and digest) responses (Lang, Bradley, & Cuthbert, 1997). This finding is suggestive of an orienting response, in which one's parasympathetic nervous system is activated to slow down and attend to new, novel, and unpredictable input (Graham, 1973; Graham & Clifton, 1966; Lacey, 1959; Lacey, J. I. & Lacey, B. C., 1974; Lang et al., 2013). Research involving affective pictures containing aversive content have reported similarly sustained decelerations of HR, or orienting responses, throughout the exposure time (Lang et al., 1997; Lang et al., 2013). It is worth noting here that previous research in SLA heralded orienting as a facilitative component for detection within the 3-component model of attention, which can be linked with resultant cognitive performance and learning (Leow, 1998; Tomlin & Villa, 1994).

Results of the analysis of HRV scores over time revealed a borderline main effect of time; however, this effect was not carried forward into the regression analyses because state-level HRV measures have yet to be established in the field of psychophysiology (Park et al., 2013; Thayer et al., 1996). Instead, I used the baseline measure from the 90-second practice clip period as the operationalization of trait-level differences of parasympathetic control of the relationship between respiration and HR (Berntson, Cacioppo, & Quigley, 1993; Grossman, 1992; Grossman & Svebak, 1987; Hugdahl, 1995; Porges, 1986).

Finally, the analysis of the change in SCL resulted in a significant main effect of time. Pairwise comparisons revealed that the change in SCL mean scores at Time 1 significantly differed from all subsequent time points across both mood induction groups; that is, both groups experienced a heightened skin conductance level during the first film followed by a steady decline over the remaining clips. The direction of the change scores was negative, which provides additional evidence of orienting and habituation, which is defined as the "ubiquitous and adaptive phenomenon whereby subjects become less responsive to familiar and non-significant stimuli" (Dawson et al., 2007, p. 167; Lang et al., 1997; Lang et al., 1993; Libby, Lacey, & Lacey, 1973; Winton, Putnam, & Krauss, 1984). Taken together, the physiological analyses confirm that the mood induction did have an impact on participants' involuntary responses to the stimuli. Now, I turn to investigate just how much impact these psychophysiological measures—in conjunction with the mood induction—had on subsequent vocabulary posttests.

The effect of the mood induction on vocabulary learning on a paired-associates learning task

When the mood induction variable was entered into the regression analyses, it significantly accounted for variation in the free recall and receptive vocabulary posttests. This finding concurs with the previous studies in SLA literature pertaining to mood or anxiety-provoking situations; generally speaking, performance declined when participants were exposed to a negative mood through film clips (Fox et al., in preparation; Miller, 2016; Miller et al., 2017), intimidated by the presence of a video camera (MacIntyre & Gardner, 1994a), or treated coldly by a brusque researcher (Steinberg & Horwitz, 1986). Additionally, existing research on the perseverative consolidation theory of learning suggests that those who are in a state of higher arousal during a learning task perform worse on immediate paired-associates recall tasks compared to those in a more neutral state (Kleinsmith & Kaplan, 1963, 1964; Kleinsmith, Kaplan, & Trate, 1963; McLean, 1969). This theory also suggests that long-term effects of arousal leads those in the higher arousal group to outperform their neutral counterparts on delayed paired-associates tasks, but such a finding does not align with the present study's immediate and delayed posttest reported in Miller et al. (2017). He identified a consistent deterioration of posttest scores over time across both neutral and negative groups.

Next, the correlation table revealed significant relationships between mood induction and the self-report delta SAM, self-report anxiety, and Mean HR measures. This suggests that mood induction, which was entered first into the hierarchical regressions, claimed much of the shared variance among these variables. This may be part of the reason why these variables were less influential in the subsequent full-model analyses (RQ3).

The effects of physiology on vocabulary learning on a paired-associates learning task (a) in isolation, and (b) when controlling for the mood induction, state-level self-report measures, and self-report measures of trait-level anxiety

When the three physiology measures were entered into the regression models, Mean HR and Delta SCL emerged as significant predictor variables in the free recall and productive tests, and the Baseline HRV measure was also significant in the productive posttest model. The directionality of the regression coefficients was positive, suggesting that those with higher Mean HR, greater Delta SCL, and higher Baseline HRV measures performed better on the vocabulary posttests. The second portion of RQ 3 investigated the role of psychophysiology on vocabulary learning when mood induction, self-report Delta SAM, and self-report Anxiety scores were first entered into the model to control for all known variation. Results revealed that the Delta SCL measure was the only significant predictor for the free recall and productive posttests. To interpret this finding, the individuals who had the larger change in skin conductance levels from the first film to the final film, those who were initially aroused but then recovered well, performed better than those with less variation in their skin conductance levels over the course of the film exposure. Research from psychophysiology has termed individual differences in the rate of SCL habituation "electrodermal lability" (Crider, 1993; Lacey & Lacey, 1958; Mundy-Castle & McKiever, 1953), in which individuals who exhibit great lability, or variability, are labeled "labiles," and those with narrower-ranging electrodermal responses are called "stabiles." Findings have revealed that labiles perform better than stabiles on tasks that demand sustained vigilance. By extension, I advocate that any aspect of second language learning, particularly vocabulary learning, requires sustained vigilance and attention. The concept of electrodermal lability aligns well with the findings of the present study: participants with greater changes in SCL levels from the beginning to the end of the film exposure performed better on the subsequent vocabulary posttests. Furthermore, researchers have noted that "lability reflects the ability to allocate information processing capacity to stimuli that are to be attended" (Dawson et al., 2007, p. 173). Thus, electrodermal lability also relates to the literature on attention and orienting, which as previously noted, is considered to be the "gateway to attention" (Lang et al., 2013, p. xix). Those with greater orienting effects are better prepared to attend to novel stimuli, which may facilitate later detection and uptake of material (Leow, 1998; Tomlin & Villa, 1994).

By far, the Delta SCL measure proved to be the most influential variable in both the physiology-only and full model hierarchical regressions. However, I would be remiss to overlook the role of the Mean HR variable, which proved to be a significant predictor in the physiology-only regressions. Upon initial consideration, it is puzzling how the manipulation checks provided evidence for the orienting effect in both the Mean HR and Delta SCL measures, yet only Delta SCL took center stage as the significant predictor variable in the subsequent full model hierarchical regressions. Upon further inspection of the initial correlations table, however, a possible solution to this conundrum surfaced. As previously noted, the correlation table revealed significant relationships between mood induction Mean HR measures ($r = -.48$, $p < .001$); the relationship between the mood induction and Delta SCL, however, was not significant ($r = -.05$, $p = .68$). Therefore, one possibility why Mean HR did not predict vocabulary learning in the full model is that it correlated too highly with mood induction. Because mood induction was entered into the full model regressions first, it already accounted for the shared variance with Mean HR in vocabulary test scores. This was not the case for the Delta SCL, however, because there was no significant relationship with mood induction.

Finally, it is notable that the two self-report factor scores, state-level Delta SAM and the trait-level anxiety scores, did not significantly account for variability in the posttests. This does not mean that some of the component scores were not informative; rather, it may be an indication that the self-report data were considerably collapsed in order to enable a greater focus on the psychophysiological measures. See Fox et al. (in preparation) for a detailed discussion of the impact of mood induction and specific trait-level anxiety factor scores on the vocabulary posttests.

In summary, we observed evidence for an orienting effect in two of the psychophysiological measures: mean HR (main effect of mood induction) and Delta SCL (main effect of time). Orienting refers to an individual's increased levels of attention and receptivity to novel input (Graham, 1973; Graham & Clifton, 1966; Lacey, 1959; Lacey, J. I. & Lacey, B. C., 1974; Lang et al., 2013). Although both variables' manipulation checks pointed to the same orienting effect, only the Delta SCL variable was predictive in two the subsequent full model hierarchical regressions of the vocabulary posttests. Also, the mood induction variable significantly contributed to two of the three posttests when initially entered as the only predictor variable. In sum, changes in emotion–represented by the Delta SCL variable–did significantly contribute to participants' performance on the vocabulary learning posttests.

Limitations

Due to the paucity of previous research on the role of psychophysiology in SLA, the literature has necessarily been sourced from other disciplines from different periods of methodological history. This expedition into the early days of research on emotion, psychophysiology, and learning has demonstrated just how far the cognitive-emotional bridge has been extended since its initial footings were placed. Most of the other limitations fall within the domain of methodology; I will begin with issues pertaining to the sample, followed by challenges of specific materials, data extraction, and scoring.

First, this project was limited by a relatively small sample size, not uncommon in the field of SLA (Plonsky, 2013, 2014). Nevertheless, the results of the separate analyses revealed satisfactory indices of power and other measures to support the decision to use inferential statistics; for example, exploratory factor analyses satisfied general benchmarks such as KMO which confirmed the appropriateness of a factor analysis for this sample (Loewen & Gonulal, 2015). It is remarkable that the results of analyses are as revealing as they are, considering the number of cases that had to be eliminated in the final regression models due to missing data, computer-based technical errors in the data collection phase, and influential outliers. Nevertheless, this serves as a lesson for the next round of research involving psychophysiological data collection: build in at least 20% extra participants in anticipation of losing data (Alhabash, 2016; Moser, 2016). The small sample size does not negate the findings of this study, but rather tempers eventual generalizations of the results to the greater population of adult second language learners.

A second issue regarding participant sampling was the less-than-random assignment of participants to the two mood induction groups. During the initial screening interview that took place over the phone with all potential participants, fewer than 5 individuals responded yes to the question, “Are you highly sensitive to violent and graphic images?” These individuals, if they fit the remaining criteria for inclusion in the study, were then assigned to the neutral film group. Unfortunately, the researchers did not tag which participants these were and whether or not some of their data was ultimately eliminated due to the aforementioned reasons. Therefore, the full effects of a negative mood may be underestimated in the present study because those most susceptible to a mood induction were assigned to the neutral comparison group. Retrospectively, in order to control as many potentially influential factors as possible, these individuals would not have been invited to participate in the study.

The next limitation pertains to the operationalization of cognition. Within SLA, there are a variety of tasks (i.e., recognition, memory, comprehension, grammatical judgment) and linguistic features that could be selected as the case study of cognition and attention (Simard & Wong, 2001). In the case of the present study, I employed a paired-associates vocabulary-learning task of 24 novel Indonesian words as an appropriate operationalization of deliberate and intentional learning inherent to tasks of vocabulary learning, memory, and cognition (de Groot & Smedinga, 2014; de Groot & van Hell, 2005; Elgort, 2011; Elgort & Warren, 2014; Schmitt, 2010a, 2010b; Simard & Wong, 2001). Paired-associates learning is a cornerstone task within assessments predicting foreign language learning success, such as the Modern Language Aptitude Test (Carroll & Sapon, 1959), Language Aptitude Battery (Pimsleur, 1966), Cognitive Ability for Novelty in Acquisition of Language as applied to foreign language test (CANAL-FT; [Grigornko, Sternberg, & Ehrman, 2000]), and the High-Level Language Aptitude Battery (Linck et al., 2013). By extension, although debate surrounds the connectedness of paired-associates learning with second language acquisition, its prevalent use in language aptitude assessments underscores its usefulness in demonstrating how language learners may perform on other language-related tasks.

Within the domain of vocabulary, another limitation was the apparent ineffectiveness of the free recall task to provide evidence for negative effect of the mood induction. I hypothesize that this difference was not statistical but rather due to the instructors by which we elicited free recall: “any English or Indonesian words remembered” by the participants. I think this broad elicitation approach elicited too broad of a construct of free recall. That is, some students may have provided more English words as a strategy that capitalized on their L1 memory, others may have supplied more Indonesian words in order to showcase their language learning, and still other participants may have provided as many English—Indonesian word pairs they remembered. All in all, the results were as mixed as the participants’ potential motivations for completing the free recall task. In future research projects employing free recall, I would advise researchers to elicit only Indonesian words for the task.

The final area of limitations concerns the psychophysiological data extraction and scoring procedures, most of which are relatively new and unfamiliar to SLA researchers. First, the use of the same time windows across all three physiological measures may have led to a misalignment of time phases under analysis. On the one hand, this was the most intuitive and straightforward procedure to apply for HR, HRV, and SCL. However, it has been reported that SCL signals occur on a much slower timeline, and because of this, many researchers expand the window of time that they analysis the SCL signal in order to capture its entirety (Dawson et al., 2007; Dawson et al., 2011). I prioritized uniformity of extraction across the three measures, but future researchers may desire to consider the nature and time course of the different psychophysiological signals.

Another challenge of utilizing psychophysiology data was the inevitable noise in the data, such as the wide psychophysiological variability across individuals (Mates & Joaquin, 2013). It is a large assumption to draw one-to-one relationships between such basic physiologic reactions and higher level constructs such as emotion, cognition and learning (Lane et al., 2000). As Dörnyei (2009) noted, there is no such thing as a monolithic language learner experience; the present study underscores the seemingly unlimited sources of variability beyond the typical individual differences in SLA research such as the following: personality traits (such as introversion vs. extraversion), time of day of the experiment, the presence of stimulant drugs, and activity levels (Revelle & Loftus, 1992; Robinson, 1995). In summary, the field of psychophysiology and SLA must account for both internal and external individual differences.

Finally, it would have been very informative to track the changes in physiology and emotional self-reports beyond the course of the emotional induction to the subsequent vocabulary learning and posttest phases. However, these two phases were not included due to the uncontrolled variability in the time marking for those two phases. That is, while most psychophysiological researchers program automatic markers for the start and end time of specific events (i.e., the start and end of the vocabulary learning trials), it would have been difficult to set such markers for the posttests because they were paper-based and researcher administered. Future projects would benefit from encoding markers into the experimental trial and using computer-based vocabulary posttests with a timer. This would optimize the time of the experiment, eliminate potential interference or variability in the experiment due to the presence of the researcher, streamline the data cleaning and processing procedures, and greatly reduce the risk of human error in manually-calculated time markers.

Pedagogical implications and future directions

Due to the exploratory nature of the current study, which was conceived as an attempt to build bridges between SLA and psychology, pedagogical implications blend into future research directions. After all, it would be unreasonable to suggest that language learners and instructors monitor their changes in SCL levels throughout the learning process. What can be recommended, however, is in regards to the type of environment conducive for language learning. The fact that participants in the negative mood induction group performed worse on subsequent vocabulary posttests underlines the importance for language instructors to provide a neutral and non-threatening space for learning to take place (Arnold, 2011; Dewaele, 2013; Gregersen & MacIntyre, 2014). Dewaele (2013, pp. 680-681) noted "the teacher should strive to create a low-threat, positive learning environment where teacher support and group solidarity will encourage anxious learners to participate and where judicious praise might promote their self-perception as FL users." Ehrman, Leaver, and Oxford (2003) offer a number of language study skill principles, including one that recommends language learners develop the ability to "manage feelings…to use more cognitive strategies, since negative feelings reduce the effectiveness of most learning activities. Appropriate self-efficacy promotes persistence in the face of difficulty" (p. 319). There is no doubt that students will be faced with internal and external stressors and threats, such as fear of negative evaluation or foreign language classroom anxiety. Nevertheless, instructors can take an active role in educating learners about affective strategies or techniques, like self-reinforcement and positive self-talk, which can assist learners in gaining more control over their emotions, attitudes, and motivations related to language learning (Oxford & Crookall, 1989).

As previously mentioned, there are also a number of technical improvements that could be made to future studies of this nature. One such improvement would be to intensify and extend the amount of time participants are exposed to the various mood inductions in order to elicit stronger and more discriminable patterns of physiology. Another improvement would be to include different types of language tasks and linguistic features to more comprehensively represent second language acquisition (Chwilla et al., 2011; Simard & Wong, 2001; Verhees et al., 2015). Furthermore, to increase the ecological validity of the project, participants could be actual language students learning a language they are already studying to increase the investment they may have. Ideally, future interdisciplinary projects would include a dedicated team of researchers who could contribute his or her specific expertise to the group's efforts. Particular competencies in online processing methodologies–such as eye tracking, functional magnetic resonance imaging (fMRI), and event-related potential (ERP) research–would undoubtedly provide a more holistic description of the mind-body link. This domain of research is promising, and there is still much to be uncovered.

Conclusion

The present study set out to construct a bridge between the fields of second language acquisition and psychology. In particular, its aim was to determine to what extent emotions impacted cognition. Emotion was operationalized in three ways: self-reported trait-level anxiety factor scores; self-reported state-level changes in valence (happiness) and arousal (excitement); and state-level physiologic changes in heart rate, heart rate variability, and skin conductance levels. One group of participants' emotional states were manipulated by exposing them to films clips of an aversive and potentially distressing nature; the comparison group was exposed to neutral film clips. All participants engaged in a paired-associates vocabulary-learning task after the mood induction, and I chose this task to represent intentional or explicit learning associated with higher-order cognitive processing. Results revealed that the negative mood induction group of participants performed worse on the vocabulary posttests; additionally, those who exhibited a greater change (recovery) in SCL score also performed better on the subsequent tests. There were no significant interactions between mood induction and other predictor variables.

These findings build on the results of the preceding studies on this same dataset (Fox & Miller, 2015; Fox et al., in preparation; Miller, 2016; Miller et al., 2017). Results from Miller et al. (2017) revealed that the negative mood induction interfered with the immediate and delayed vocabulary posttest scores; that is, those in the negative mood induction group consistently performed worse than their neutral condition counterparts. Furthermore, there were no significant group (negative or neutral) or time differences in participants' working-memory capacity tests pre- and post-mood induction. The null results for the WMC tests rule out a cognitive variable (i.e., working memory) as the underlying agent of change triggered by the mood induction.

While Miller et al. (2017) ruled out working memory as a mediating variable, the results of the present study suggest that Delta SCL and Mean HR may be mediators. The fact that psychophysiological measures predicted vocabulary learning in the paired-associates task aligns well with the results from Fox et al. (in preparation), in which the role of the anxiety questionnaires, currently subsumed under the trait-level self-report anxiety factor score, were investigated for their predictive ability for vocabulary learning. Specifically, participants who reported higher than average STICSA somatic distress (i.e., physical symptoms of anxiety, such as sweaty palms, butterflies in the stomach, and a racing heartbeat) outperformed their average or below average anxious counterparts in the baseline (neutral) condition. However, a negative mood induction hampered vocabulary learning and did so disproportionately for more anxious individuals. Thus, the fact that the self-reported, trait level STICSA somatic distress factor score–in interaction with mood induction–positively predicted better performance on the vocabulary posttests in the neutral condition suggested some connection between the body and cognition, as measured by the paired-associates vocabulary learning task.

The results of the present study, therefore, validate the suspicion that somatic distress does have an impact on vocabulary performance on a paired-associates task. The Delta SCL, the most direct indicator of arousal, was a significant predictor of vocabulary learning across both conditions. While not immediately applicable to language learners and instructors, the importance of heightened attention and orienting (represented by higher Mean HR measures) and emotional recovery (represented by large Delta SCL measures) proved to facilitate vocabulary learning in the present study.

The aim of the bridge-building enterprise was to establish a link between the fields of SLA and psychology through an investigation of the intersection of emotion and cognition. Specifically, the overarching goal of the present study was to examine how emotions—represented by self-report, psychophysiological variables, and responses to the mood induction—influenced cognition—operationalized by a paired-associates vocabulary-learning task. I did uncover evidence that both the psychophysiological measures of Mean HR and Delta SCL did influence subsequent vocabulary learning. Emotion, in its purest, most basic sense, did have an impact on a higher-order cognitive task. Stepping away from the details of the bridge's deck, hangers, cables, and towers, I realize that the present study has accomplished what it set out to do: to add another link in the bridge between SLA and psychology as well as between emotion and cognition.

APPENDICES

Appendix A

Self-Assessment Manikin (SAM)

Please rate how you feel at this exact moment on the following scales:

Sad
Unpleasant

Happy
Pleasant

Relaxed
Unaroused

Stimulated
Aroused

Controlled
Submissive

Controlling
Dominant

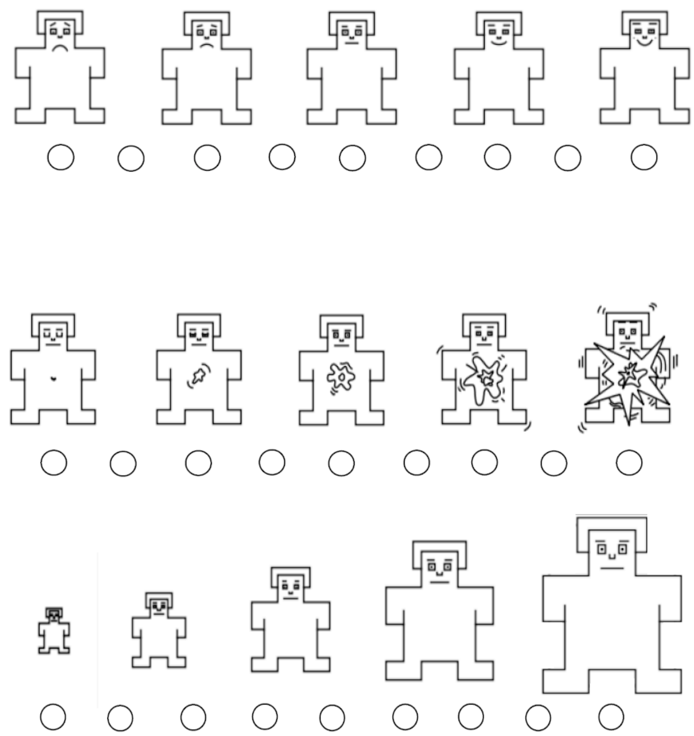

Figure 15. Self-Assessment Manikin (SAM)

Appendix B

Target Indonesian words with pictorial representations and number of letters

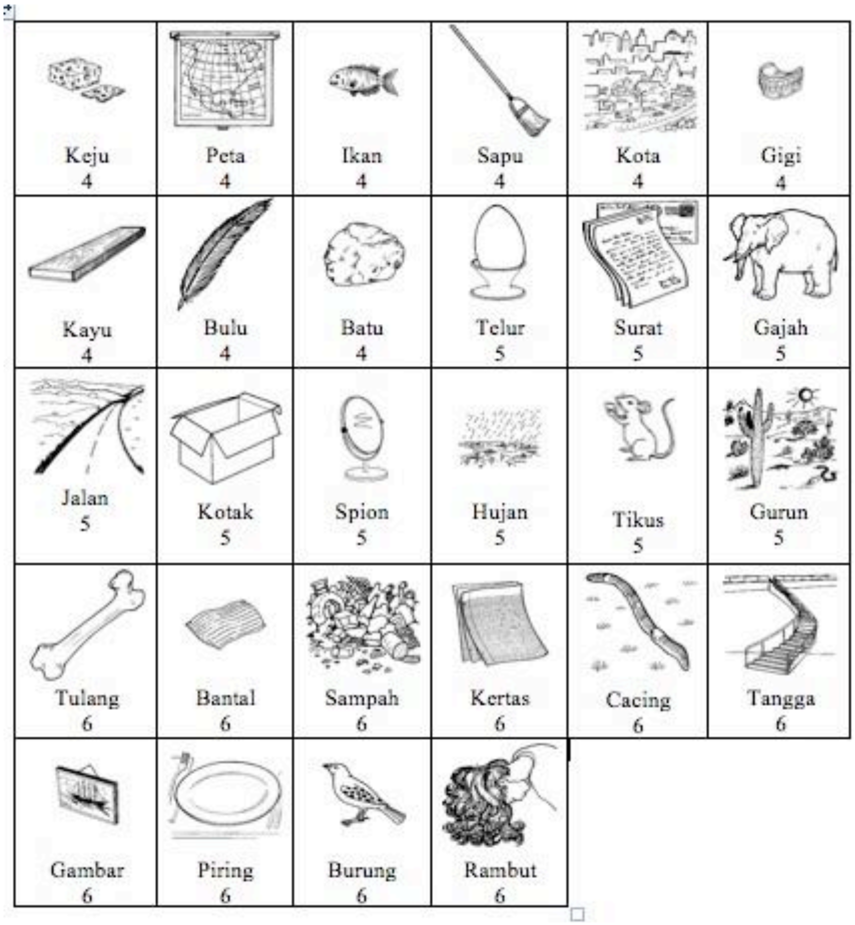

Figure 16. Target Indonesian words

Appendix C

Sample lexical training item

Burung = Bird

Figure 17. Sample lexical training item

Appendix D

Instructions for vocabulary posttests

Free recall: This test is based on the words you have just learned. Write as many words as you remember. You will have 2 minutes and 24 seconds to complete this test. Let the experimenter know if you have any questions. Good luck!

Productive: This test is based on the words you have just learned. Write the correct translation for the English word provided. You will have 2 minutes and 24 seconds to complete this test. Let the experimenter know if you have any questions. Good luck!

Receptive: This test is based on the words you have just learned. Write the correct English translation for the word provided. You will have 2 minutes and 24 seconds to complete this test. Let the experimenter know if you have any questions. Good luck!

Appendix E

Lexical posttest scoring rubric

Table 28

Lexical posttest scoring rubric (Barcroft, 2003)

Points	*Description*
0	Nothing is written; letters present do not meet criteria for .25 point; English word is written.
0.25	Any 1 letter is correct; 25-49% letters present; word contains the correct amount of syllables.
0.50	50% of the word is written, including 25-49% letters correct and 50-74% letters present.
0.75	75% of the word is written, including 50-99% letters correct and 75-99% letters present; 100% letters correct but other letters added.
1.0	Entire word is written, including 100% letters correct with proper accent marks.

(1) "Correct" refers to any letter written and placed in its correct position within a word; "present" refers to any letter written but not placed in its correct position.

(2) Determin percentages by dividing letters correct and letters present by the number of letters in the target word. If more letters are written than are in the target word, divide by the larger number.

(3) If the same target word is written more than once, score it only once in the space where it should be written or, if it is not written in the correct space, score it in the first space where it is written based upon the target word for that space.

Appendix F

Foreign Language Classroom Anxiety Scale (FLCAS)

<u>Instructions:</u>
Read each statement below and respond AS THEY APPLY TO YOUR STUDY OF A FOREIGN LANGUAGE. Please indicate your responses on the scale below using the following criteria:

Circle 1 if you Strongly disagree with the statement
Circle 2 if you Disagree with the statement
Circle 3 if you Neither disagree nor agree with the statement
Circle 4 if you Agree with the statement
Circle 5 if you Strongly agree with the statement

	Strongly Disagree	Disagree	Neither disagree nor agree	Agree	Strongly Agree
1. I never feel quite sure of myself when I am speaking in my foreign language class.	1	2	3	4	5
2. I don't worry about making mistakes in class.	1	2	3	4	5
3. I tremble when I know that I'm going to be called on in language class.	1	2	3	4	5
4. It frightens me when I don't understand what the teacher is saying in the foreign language.	1	2	3	4	5
5. It wouldn't bother me at all to take more foreign language classes.	1	2	3	4	5
6. During language class, I find myself thinking about things that have nothing to do with the course.	1	2	3	4	5
7. I keep thinking that the other students are better at languages than I am.	1	2	3	4	5
8. I am usually at ease during tests in my language class.	1	2	3	4	5
9. I start to panic when I have to speak without preparation in language class.	1	2	3	4	5
10. I worry about the consequences of failing my foreign language class.	1	2	3	4	5
11. I don't understand why some people get so upset over foreign language classes.	1	2	3	4	5
12. In language class, I can get so nervous I forget things I know.	1	2	3	4	5
13. It embarrasses me to volunteer answers in my language class.	1	2	3	4	5
14. I would not be nervous speaking the foreign language with native speakers.	1	2	3	4	5

15. I get upset when I don't understand what the teacher is correcting.	1	2	3	4	5
16. Even if I am well prepared for language class, I feel anxious about it.	1	2	3	4	5
17. I often feel like not going to my language class.	1	2	3	4	5
18. I feel confident when I speak in foreign language class.	1	2	3	4	5
19. I am afraid that my language teacher is ready to correct every mistake I make.	1	2	3	4	5
20. I can feel my heart pounding when I'm going to be called on in language class.	1	2	3	4	5
21. The more I study for a language test, the more confused I get.	1	2	3	4	5
22. I don't feel pressure to prepare very well for language class.	1	2	3	4	5
23. I always feel that the other students speak the foreign language better than I do.	1	2	3	4	5
24. I feel very self-conscious about speaking the foreign language in front of other students.	1	2	3	4	5
25. Language class moves so quickly I worry about getting left behind.	1	2	3	4	5
26. I feel more tense and nervous in my language class than in my other classes.	1	2	3	4	5
27. I get nervous and confused when I am speaking in my language class.	1	2	3	4	5
28. When I'm on my way to language class, I feel very sure and relaxed.	1	2	3	4	5
29. I get nervous when I don't understand every word the language teacher says.	1	2	3	4	5
30. I feel overwhelmed by the number of rules you have to learn to speak a foreign language.	1	2	3	4	5
31. I am afraid that the other students will laugh at me when I speak the foreign language.	1	2	3	4	5
32. I would probably feel comfortable around native speakers of the foreign language.	1	2	3	4	5
33. I get nervous when the language teacher asks questions which I haven't prepared in advance.	1	2	3	4	5
Comment Box: Do you have any questions or comments for the researcher?					

Appendix G

Mood and Anxiety Symptom Questionnaire (MASQ)

<u>Instructions:</u>

Below is a list of feelings, sensations, problems, and experiences that people sometimes have. Read each item and then circle the appropriate number next to each statement. Use the choice that best describes *<u>how much</u>* you have felt or experienced things this way *<u>during the past week, including today</u>*. Use the following scale when answering each item.

Circle 1 for Not at all
Circle 2 for A little bit
Circle 3 for Moderately
Circle 4 for Quite a bit
Circle 5 for Extremely

	Not at all	**A little bit**	**Moderately**	**Quite a bit**	**Extremely**
1. Startled easily	1	2	3	4	5
2. Felt cheerful	1	2	3	4	5
3. Hands were shaky	1	2	3	4	5
4. Felt optimistic	1	2	3	4	5
5. Felt really happy	1	2	3	4	5
6. Was short of breath	1	2	3	4	5
7. Was proud of myself	1	2	3	4	5
8. Felt faint	1	2	3	4	5
9. Felt unattractive	1	2	3	4	5
10. Had hot or cold spells	1	2	3	4	5
11. Felt like I was having a lot of fun	1	2	3	4	5
12. Hands were cold or sweaty	1	2	3	4	5
13. Felt withdrawn from other people	1	2	3	4	5
14. Felt like I had a lot of energy	1	2	3	4	5
15. Was trembling or shaking	1	2	3	4	5
16. Had trouble swallowing	1	2	3	4	5
17. Felt really slowed down	1	2	3	4	5
18. Felt dizzy or lightheaded	1	2	3	4	5
19. Felt really "up" or lively	1	2	3	4	5
20. Had pain in my chest	1	2	3	4	5
21. Felt really bored	1	2	3	4	5
22. Felt like I was choking	1	2	3	4	5

23. Looked forward to things with enjoyment	1	2	3	4	5
24. Muscles twitched or trembled	1	2	3	4	5
25. Had a very dry mouth	1	2	3	4	5
26. Felt like I had a lot of interesting things to do	1	2	3	4	5
27. Was afraid I was going to die	1	2	3	4	5
28. Felt like I had accomplished a lot	1	2	3	4	5
29. Felt like it took extra effort to get started	1	2	3	4	5
30. Felt like nothing was very enjoyable	1	2	3	4	5
31. Heart was racing or pounding	1	2	3	4	5
32. Felt like I had a lot to look forward to	1	2	3	4	5
33. Felt numbness or tingling in my body	1	2	3	4	5
34. Felt hopeful about the future	1	2	3	4	5
35. Felt like there wasn't anything interesting or fun to do	1	2	3	4	5
36. Seemed to move quickly and easily	1	2	3	4	5
37. Felt really good about myself	1	2	3	4	5
38. Had to urinate frequently	1	2	3	4	5

Comment Box: Do you have any questions or comments for the researcher?

Appendix H

Penn State Worry Questionnaire (PSWQ)

<u>Instructions:</u>
Rate each of the following statements on a scale of "1" (not at all typically of me) to "5" (very typical of me). Please do not leave any items blank.

Not at all typical of me			**Very typical of me**	
1	**2**	**3**	**4**	**5**

1.	If I do not have enough time to do everything, I do not worry about it.	1	2	3	4	5
2.	My worries overwhelm me.	1	2	3	4	5
3.	I do not tend to worry about things.	1	2	3	4	5
4.	Many situations make me worry.	1	2	3	4	5
5.	I know I should not worry about things, but I just cannot help it.	1	2	3	4	5
6.	When I am under pressure I worry a lot.	1	2	3	4	5
7.	I am always worrying about something.	1	2	3	4	5
8.	I find it easy to dismiss worrisome thoughts.	1	2	3	4	5
9.	As soon as I finish one task, I start to worry about everything else I have to do.	1	2	3	4	5
10.	I never worry about anything.	1	2	3	4	5
11.	When there is nothing more I can do about a concern, I do not worry about it any more.	1	2	3	4	5
12.	I have been a worrier all my life.	1	2	3	4	5
13.	I notice that I have been worrying about things.	1	2	3	4	5
14.	Once I start worrying, I cannot stop.	1	2	3	4	5
15.	I worry all the time.	1	2	3	4	5
16.	I worry about projects until they are done.	1	2	3	4	5

Appendix I

State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA)

<u>Instructions:</u>
Below is a list of statements which can be used to describe how people feel. Please read each statement carefully and respond HOW OFTEN, IN GENERAL, THE STATEMENT IS TRUE OF YOU. Indicate your responses on the scale below using the following criteria:

Circle 1 if *Not at all*
Circle 2 if *A little*
Circle 3 if *Moderately*
Circle 4 if *Very much so*

	Not at all	**A little**	**Moderately**	
1. My heart beats fast.	1	2	3	4
2. My muscles are tense.	1	2	3	4
3. I feel agonized over my problems.	1	2	3	4
4. I think that others won't approve of me.	1	2	3	4
5. I feel like I'm missing out on things because I can't make up my mind soon enough.	1	2	3	4
6. I feel dizzy.	1	2	3	4
7. My muscles feel weak.	1	2	3	4
8. I feel trembly and shaky.	1	2	3	4
9. I picture some future misfortune.	1	2	3	4
10. I can't get some thought out of my mind.	1	2	3	4
11. I have trouble remembering things.	1	2	3	4
12. My face feels hot.	1	2	3	4
13. I think that the worst will happen.	1	2	3	4
14. My arms and legs feel stiff.	1	2	3	4
15. My throat feels dry.	1	2	3	4
16. I keep busy to avoid uncomfortable thoughts.	1	2	3	4
17. I cannot concentrate without irrelevant thoughts intruding.	1	2	3	4
18. My breathing is fast and shallow.	1	2	3	4
19. I worry that I cannot control my thoughts as well as I would like to.	1	2	3	4
20. I have butterflies in the stomach.	1	2	3	4
21. My palms feel clammy.	1	2	3	4

Appendix J

Variable operationalizations

Table 29

Variable operationalizations

Variable Name	**Definition**	**Unit of Measurement**	**Variable type**
Mood induction	This is the group to which the participant is assigned according to which films the participant was exposed to	Negative Neutral	Independent Categorical
Time	The periods in the experiment of Day 1 that are significant to the analysis, labeled ordinally, from the beginning of the study to the end	Time 1=initial Time 2=immediately following the film exposure Time 3=end of day 1 procedure	Independent Ordinal
Self-assessment manikin (SAM)	Emotional self-report used periodically during the experiment	Three Likert scales (1 to 9); valence and arousal	Independent Ordinal
Δ SAM	Change in Emotional self-report from Time 1 to 2	Three Likert scales (1 to 9); valence and arousal	Independent Continuous
Anxiety factor scores	A participant's self-reported level of anxiety from Fox (2016): FLCAS 1- foreign language anxiety FLCAS 2-low engagement MASQ1- positive emotionality MASQ2- somatic distress PSWQ- worry STICSA1- cognitive distress STICSA 2- somatic distress	z-scores, means of 0	Independent Continuous
Heart rate (HR)	Each of the 6 films resulted in a baseline-corrected (active –	IBI—interbeat intervals, in milliseconds	Independent Continuous

	ITBL) mean score over the course of each film (40sec.). Each 10-second interlude between the films is used as an inter-active phase correction for baseline. I used this measure to operationalize state heart rate.	BPM—beats per minute	
Baseline Heart rate variability (HRV_BL)	This is the measure of heart rate variability during the 90 seconds of practice film clips. I used this measure to operationalize trait heart rate variability.	High frequency respiratory sinus arrhythmia (HF RSA) is a metric of the frequency domain of HRV. ms^2 –milliseconds squared	Independent Continuous
Skin conductance level (SCL)	This is the average of 6 mean scores over the course of the active films (45sec.= 3 sec before active film clip, 40 sec film clip, and 2 seconds after end of active film clip). I used this measure to operationalize state skin conductance level.	Microsiemens (µs)	Independent Continuous
Free recall (FR)	This is the first vocabulary posttest in which participants must write any words (English or Indonesian) that they remember from the learning phase.	Out of 24 points (initially 48, because 24 English and Indonesian, but we divided it by two in order to have a comparable scale with other two measures). Scoring rubric follows Barcroft (2003).	Dependent Continuous
Productive (P1)	This is the second vocabulary posttest in which participants must translate English words into Indonesian words they just learned.	Out of 24 points. Scoring rubric follows Barcroft (2003).	Dependent Continuous
Receptive (R1)	This is the third vocabulary posttest in which participants must translate Indonesian words into English.	Out of 24 (Note: this measure violates the regression assumption of unbounded variables with an upper-bounded negative skew. In Fox (2016), I reversed scores (x-24), then log transformed data, had to use the zero-inflated negative binomial regression to analyze data and account for perfect (0) scores).	Dependent Continuous

REFERENCES

REFERENCES

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral Cognitive Neuroscience Review, 1*(1), 21-62.

Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquart (Eds.), *Reading in a Foreign Language* (pp. 1-21). New York: Longman.

Alhabash, S. (2016). [Data collection of electrodermal activity measures].

Alpert, R., & Haber, R. N. (1960). Anxiety in academic achievement situations. *Journal of Abnormal Social Psychology, 61*(2), 207-215.

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behav Res Methods Instrum Comput, 31*(4), 578-602.

Altarriba, J., & Canary, T. M. (2004). The Influence of Emotional Arousal on Affective Priming in Monolingual and Bilingual Speakers. *Journal of Multilingual and Multicultural Development, 25*(2-3), 248-265. doi:10.1080/01434630408666531

Amiel, H. F. (2010). [A blessing].

Amiez, C., Procyk, E., Honore, J., Sequeira, H., & Joseph, J. P. (2003). Reward anticipation, cognition, and electrodermal activity in the conditioned monkey. *Experimental Brain Research, 149*(3), 267-275. doi:10.1007/s00221-002-1353-9

Analyzer, B. V. (2006). User Manual. 2006. *Brain Products GmbH*.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.

Appleby, R. (2010). *ELT, gender and international development*. Brookvale: Multilingual Matters.

Arnaiz, P., & Guillén, F. (2012). Foreign language anxiety in a Spanish university setting: Interpersonal differences. *Revista de Psicodidáctica, 17*(1), 5-26.

Arnold, J. (2011). Attention to affect in language learning. *Anglistik: International Journal of English Studies, 22*, 11-22.

Association, A. H. (2015). Target heart rates. Retrieved from http://www.heart.org/HEARTORG/HealthyLiving/PhysicalActivity/FitnessBasics/Target-Heart-Rates_UCM_434341_Article.jsp - .WKp-uxIrLox

Astika, G. G. (1993). Analytical assessments of foreign students' writing. *RELC Journal, 24*(1), 61-70.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189-208.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review, 105*(1), 158.

Bailey, K. M. (1996). *Voices from the language classroom: Qualitative research in second language education*. Cambridge: Cambridge University Press.

Baldaro, B., Mazzetti, M., Codispoti, M., Tuozzi, G., Bolzani, R., & Trombini, G. (2001). Autonomic reactivity during viewing of an unpleasant film. *Perceptive Motor Skills, 93*(3), 797-805. doi:10.2466/pms.2001.93.3.797

Barcroft, J. (2003). Distinctiveness and bidirectional effects in input enhancement for vocabulary learning. *Applied Language Learning, 13*(2), 133-159.

Barry, R. J., & Sokolov, E. N. (1993). Habituation of phasic and tonic components of the orienting reflex. *International Journal of Psychophysiology, 15*(1), 39-42.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science, 2*(4), 396-403.

Beauchaine, T. P., Gatzke-Kopp, L., & Mead, H. K. (2007). Polyvagal Theory and developmental psychopathology: emotion dysregulation and conduct problems from preschool to adolescence. *Biological psychology, 74*(2), 174-184. doi:10.1016/j.biopsycho.2005.08.008

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*(4), 506-521.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of neuroscience methods, 190*(1), 80-91. doi:10.1016/j.jneumeth.2010.04.028

Benesch, S. (2013). *Considering emotions in critical English language teaching: Theories and praxis*. New York: Routledge.

Berntson, G., Quigley, K., & Lozano, D. (2007). Cardiovascular psychophysiology *Handbook of psychophysiology* (Vol. 3, pp. 182-210). Cambridge: Cambridge University Press.

Berntson, G. G., Bigger, J. T., Jr., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., . . . van der Molen, M. W. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology, 34*(6), 623-648.

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Respiratory sinus arrhythmia: Autonomic origins, physiological mechanisms, and psychophysiological implications. *Psychophysiology, 30*(2), 183-196.

Berntson, G. G., & Stowell, J. R. (1998). ECG artifacts and heart period variability: Don't miss a beat! *Psychophysiology, 35*(1), 127-132.

BioSemi, B. (2011). BioSemi ActiveTwo.[EEG system]. *Amsterdam: BioSemi*.

Böhm, C. J. (2006). *Brief Chronicle of the Unrest in the Moluccas: 1999-2006*. Ambon, Indonesia

Bolls, P. D., Lang, A., & Potter, R. F. (2001). The effects of message valence and listener arousal on attention, memory, and facial muscular responses to radio advertisements. *Communication Research Reports, 28*(5), 627-651.

Boucher, J., & Osgood, C. E. (1969). The pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior, 8*(1), 1-8.

Boucsein, W. (1992). *Electrodermal activity*. New York: Plenum Press.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., roth, W. T., Dawson, M. E., . . . Society for Psychophysiological Research Ad Hoc Committee on Electrodermal, M. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology, 49*(8), 1017-1034. doi:10.1111/j.1469-8986.2012.01384.x

Bown, J. (2009). Self-Regulatory strategies and agency in self-instructed language learning: A situated view. *The Modern Language Journal, 93*(4), 570-583.

Bradley, M. M., & Lang, P. (2000). Measuring emotion: Behavior, feeling, and physiology. *Cognitive neuroscience of emotion, 25*, 49-59.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavioral Theraoy and Experimental Psychiatry, 25*(1), 49-59.

Brand, A. G. (1987). Hot cognition: Emotions and writing behavior. *Journal of Advanced Composition, 6*(1), 5-15.

Brosschot, J., Van Dijk, E., & Thayer, J. (2007). Daily worry is related to low heart rate variability during waking and the subsequent nocturnal sleep period. . *International Journal of Psychophysiology, 63*(1), 39-47.

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist, 45*(1), 16-28.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (Eds.). (2007). *Handbook of psychophysiology* (3 ed.). New York: Cambridge University Press.

Cannon, W. B. (1929). *Bodily changes in pain, fear, hunger, and rage*. New York: Appleton.

Cardiology, T. F. o. t. E. S. o. (1996). Heart rate variability standards of measurement, physiological interpretation, and clinical use. *European Heart Journal, 17*, 354-381.

Carroll, J., & Sapon, S. (1959). Modern Language Aptitude Test (MLAT): Manual. San Antonio, TX: Psychological Corp. Republished by Second Language Testing. *Inc., www. 2LTI. com*.

Carvalho, S., Leite, J., Galdo-Álvarez, S., & Gonçalves, Ó. F. (2012). The emotional movies database (EMDB): A self-report and psychophysiological study. *Applied Psychophysiological Biofeedback, 37*(4), 279-294.

Castaneda, A., Palermo, D. S., & McCandless, B. R. (1956). Complex learning and performance as a function of anxiety in children and task difficulty. *Child development*, 327-332.

Christianson, S.-A. (2014). *The handbook of emotion and memory: Research and theory*. New York: Psychology Press.

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.

Chwilla, D. J., Virgillito, D., & Vissers, C. T. (2011). The relationship of language and emotion: N400 support for an embodied view of language comprehension. *Journal of cognitive neuroscience, 23*(9), 2400-2414. doi:10.1162/jocn.2010.21578

Citi, L., Brown, E. N., & Barbieri, R. (2012). A real-time automated point-process method for the detection and correction of erroneous and ectopic heartbeats. *IEEE Transactions on Biomedical Engineering, 59*(10), 2828-2837. doi:10.1109/TBME.2012.2211356

Clark, M. S., & Fiske, S. T. (1982). *Affect and cognition: The seventeenth annual Carnegie symposium on cognition* (Vol. 17). New York: Psychology Press.

Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. *Second language reading and vocabulary learning*, 217-228.

Codispoti, M., Surcinelli, P., & Baldaro, B. (2008). Watching emotional movies: Affective reactions and gender differences. *International Journal of Psychophysiology, 69*(2), 90-95. doi:10.1016/j.ijpsycho.2008.03.004

Cook, E., & Turpin, G. (1997). Differentiating orienting, startle, and defense responses: The role of affect and its implications for psychopathology. *Attention and orienting: Sensory and motivational processes*, 137-164.

Corp, I. (2014). IBM SPSS Statistics for Windows (Version 23.0). Armonk, NY: IBM Corp.

Cowley, B., Fantato, M., Jennett, C., Ruskov, M., & Ravaja, N. (2014). Learning When Serious: Psychophysiological evaluation of a Technology-Enhanced Learning game. *Journal of Educational Technology & Society, 17*(1), in press.

Cowley, B., Filetti, M., Lukander, K., Torniainen, J., Henelius, A., Ahonen, L., . . . Huotilainen, M. (2016). The Psychophysiology Primer: A guide to methods and a broad review with a focus on human computer interaction. *Foundations and Trends in Human-Computer Interaction, 9*(3-4), 150-307.

Crawley, M. (2007). *The R book*. Chichester, UK: John Wiley & Sons.

Crider, A. (1993). Electrodermal response lability-stability: Individual difference correlates *Progress in electrodermal research* (pp. 173-186). New York: Springer.

Croft, R. J., Gonsalvez, C. J., Gander, J., Lechem, L., & Barry, R. J. (2004). Differential relations between heart rate and skin conductance, and public speaking anxiety. *Journal of behavior therapy and experimental psychiatry, 35*(3), 259-271. doi:10.1016/j.jbtep.2004.04.012

Cubukcu, F. (2008). A study on the correlation between self-efficacy and foreign language learning anxiety. *Journal of Theory and Practice in Education, 4*(1), 148-158.

Davis, F. B. (1944). The interpretation of frequency ratings obtained from "The Teachers Word Book.". *Journal of Educational Psychology, 35*(3), 169.

Dawson, M. E., Schell, A., & Filion, D. (2007). Chapter 7: The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 159). New York: Cambridge University Press.

Dawson, M. E., Schell, A. M., & Courtney, C. G. (2011). The skin conductance response, anticipation, and decision-making. *Journal of Neuroscience, Psychology, and Economics, 4*(2), 111.

de Groot, A., & Smedinga, H. E. (2014). Let the music play! *Studies in Second Language Acquisition, 36*(4), 681-707.

de Groot, A. M., & van Hell, A. (2005). The learning of foreign language vocabulary. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 9). Oxford: Oxford University Press.

Deane, S. (1979). *Ambon, island of spices*. London: J. Murray.

Delaney, Y. A. (2012). Research on mentoring language teachers: Its role in language education. *Foreign Language Annals, 45*(s1), s184-s202.

Demaree, H., Robinson, J., Everhart, D., & Schmeichel, B. (2004). Resting RSA is associated with natural and self-regulated responses to negative emotional stimuli. *Brain and Cognition, 56*(1), 14-23.

Demaree, H., Robinson, J., Pu, J., & Allen, J. (2006). Strategies actually employed during response-focused emotion regulation research: Affective and physiological consequences. *Cognition and Emotion, 20*(8), 1248-1260.

Dewaele, J.-M. (2002). Psychological and sociodemographic correlates of communicative anxiety in L2 and L3 production. *International Journal of Bilingualism, 6*(1), 23-38.

Dewaele, J.-M. (2004a). Blistering barnacles! What language do multilinguals swear in? *Estudios de Sociolinguistica, 5*(1), 83-105.

Dewaele, J.-M. (2004b). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development, 25*(2-3), 204-222.

Dewaele, J.-M. (2005a). The effect of type of acquisition context on perception and self-reported use of swearwords in the L2, L3, L4 and L5. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 531-559). Berlin: De Gruyter.

Dewaele, J.-M. (2005b). Investigating the psychological and emotional dimensions in instructed language learning: Obstacles and possibilities. *The Modern Language Journal, 89*(3), 367-380.

Dewaele, J.-M. (2007). The effect of multilingualism, sociobiographical, and situational factors on communicative anxiety and foreign language anxiety of mature language learners. *International Journal of Bilingualism, 11*(4), 391-409.

Dewaele, J.-M. (2013). The link between foreign language classroom anxiety and psychoticism, extraversion, and neuroticism among adult Bi- and multilinguals. *The Modern Language Journal, 97*(3), 670-684.

Dewaele, J. M., & Pavlenko, A. (2002). Emotion vocabulary in interlanguage. *Language Learning, 52*(2), 263-322.

Dewaele, J. M., Petrides, K. V., & Furnham, A. (2008). Effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: A review and empirical investigation. *Language Learning, 58*(4), 911-960.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation, 14*(20), 1-11.

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.

Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition, 13*(04), 431-469.

Douglas Fir Group, T. (2016). A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal, 100*(supp), 19-47.

Edelberg, R. (1972). Electrodermal recovery rate, goal-orientation, and aversion. *Psychophysiology, 9*(5), 512-520.

Ehrman, M. E., Leaver, B. L., & Oxford, R. L. (2003). A brief overview of individual differences in second language learning. *System, 31*(3), 313-330.

Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *The Modern Language Journal, 79*(1), 67-89.

Eilola, T. M., Havelka, J., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition & Emotion, 21*(5), 1064-1076. doi:10.1080/02699930601054109

Ekman, P. (1984). Expression and the nature of emotion. *Approaches to Emotion, 3*, 19-344.

Ekman, P., & Davidson, R. (1994). *The nature of emotion: Fundamental questions*. Oxford: Oxford University Press.

Ekman, P., Levenson, R. W., & Friesen, W. (1983). Autonomic nervous system activity distinguishes among emotions. *Science, 221*(4616), 1208-1210.

El-Sheikh, M., Hinnant, J. B., & Erath, S. (2011). Developmental trajectories of delinquency symptoms in childhood: The role of marital conflict and autonomic nervous system activity. *Journal of Abnormal Psychology, 120*(1), 16-32. doi:10.1037/a0020626

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning, 61*(2), 367-413.

Elgort, I., & Nation, P. (2010). Vocabulary learning in a second language: Familiar answers to new questions. In P. Seedhouse, S. Walsh, & C. Jenks (Eds.), *Conceptualising 'learning' in applied linguistics* (pp. 89-104). London: Palgrave Macmillan UK.

Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning, 64*(2), 365-414.

Ellis, N., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *Journal of Experimental Psychology, 46*(3), 533-558.

Engber, C. (1993). *The relationship of lexis to quality in L2 compositions*. Paper presented at the TESOL, Atlanta, GA.

Etzel, J. A., Johnsen, E. L., Dickerson, J., Tranel, D., & Adolphs, R. (2006). Cardiovascular and respiratory responses during musical mood induction. *International Journal of Psychophysiology, 61*(1), 57-69. doi:10.1016/j.ijpsycho.2005.10.025

Eysenck, M. (2014). *Anxiety and cognition: A unified theory*. New York: Psychology Press.

Eysenck, M. W. (1979). Anxiety, learning, and memory: A reconceptualization. *Journal of research in personality, 13*(4), 363-385.

Fabes, R. A., & Eisenberg, N. (1997). Regulatory control and adults' stress-related responses to daily life events. *Journal of personality and social psychology, 73*(5), 1107-1117.

Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory factor analysis*. Oxford: Oxford University Press.

Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology, 5*, 335. doi:10.3389/fpsyg.2014.00335

Fernandez, C., Pascual, J. C., Soler, J., Elices, M., Portella, M. J., & Fernandez-Abascal, E. (2012). Physiological responses induced by emotion-eliciting films. *Applied Psychophysiology and Biofeedback, 37*(2), 73-79. doi:10.1007/s10484-012-9180-7

Field, A. (2009). *Discovering statistics using SPSS* (3 ed.). Thousand Oaks, CA: Sage.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, CA: Sage.

Folse, K. S. (2004). *Vocabulary myths: Applying second language research to classroom teaching*. Ann Arbor, MI: University of Michigan Press.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & cognition, 43*(2), 226-236.

Fowles, D. C., Christie, M. J., Edelberg, R., Grings, W. W., Lykken, D. T., & Venables, P. H. (1981). Committee report. Publication recommendations for electrodermal measurements. *Psychophysiology, 18*(3), 232-239.

Fox, J. (2014a). On the car ferry.

Fox, J. (2014b). Unfinished bridge.

Fox, J. (2016). *An investigation of anxiety, hot cognition, and lexical acquisition*. Second Language Studies Program Qualifying Research Paper. Michigan State University.

Fox, J., & Miller, Z. (2015). *Playing with fire: The effects of hot cognition on emotional self-reports and lexical acquisition*. Paper presented at the Second Language Research Forum, Atlanta, GA.

Fox, J., Miller, Z., Godfroid, A., & Moser, J. (in preparation). *Emotions in Second/Adult Language Learning: An investigation of intentional vocabulary learning with physiology and self-report data.* (Dissertation), Michigan State University.

Frazier, T., Strauss, M., & Steinhauer, S. (2004). Respiratory sinus arrhythmia as an index of emotional response in young adults. *Psychophysiology, 41*(1), 75-83.

Fresco, D. M., Heimberg, R. G., Mennin, D. S., & Turk, C. L. (2002). Confirmatory factor analysis of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 40*(3), 313-323.

Grace-Martin, K. (2008). Can Likert scale data ever be continuous. *Article Alley*.

Graham, F. K. (1973). Habituation and dishabituation of responses innervated by the autonomic nervous system. *Habituation, 1*, 163-218.

Graham, F. K., & Clifton, R. K. (1966). Heart-rate change as a component of the orienting response. *Psychological bulletin, 65*(5), 305.

Green, D., & Meara, P. (1995). CALL and vocabulary teaching. *Computer Assisted Language Learning, 8*(2), 91-101.

Gregersen, T., & MacIntyre, P. (2014). *Capitalizing on language learners' individuality: From premise to practice*. Bristol: Multilingual Matters.

Gregersen, T., Macintyre, P., & Meza, M. D. (2014). The motion of emotion: Idiodynamic case studies of learners' foreign language anxiety. *The Modern Language Journal, 98*(2), 574-588. doi:10.1111/modl.12084

Grigornko, E. L., Sternberg, R. J., & Ehrman, M. E. (2000). A Theory-Based Approach to the Measurement of Foreign Language Learning Ability: The Canal-F Theory and Test. *The Modern Language Journal, 84*(3), 390-405.

Gros, D. F., Antony, M. M., Simms, L. J., & McCabe, R. E. (2007). Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-Trait Anxiety Inventory (STAI). *Psychological Assessment, 19*(4), 369-381. doi:10.1037/1040-3590.19.4.369

Gross, J. (1998). Antecedent-and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of personality and social psychology, 74*(1), 224.

Gross, J. (1999). Emotion regulation: Past, present, future. *Cognition & Emotion, 13*(5), 551-573.

Gross, J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & Emotion, 9*(1), 87-108.

Gross, J., & Muñoz, R. (1995). Emotion regulation and mental health. *Clinical Psychology: Science and Practice, 2*(2), 151-164.

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: physiology, self-report, and expressive behavior. *Journal of personality and social psychology, 64*(6), 970-986.

Grossman, P. (1992). Respiratory and cardiac rhythms as windows to central and autonomic biobehavioral regulation: Selection of window frames, keeping the panes clean and viewing the neural topography. *Biological psychology, 34*(2), 131-161.

Grossman, P., & Svebak, S. (1987). Respiratory sinus arrhythmia as an index of parasympathetic cardiac control during active coping. *Psychophysiology, 24*(2), 228-235.

Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological psychology, 74*(2), 263-285. doi:10.1016/j.biopsycho.2005.11.014

Hackley, S. A., Boelhouwer, A., Lang, P., Simons, R., & Balaban, M. (1997). The more or less startling effects of weak prestimulation-Revisted: Prepulse modulation of multicomponent blink reflexes. *Attention and orienting*, 205-227.

Hansen, A. L., Johnsen, B. H., & Thayer, J. F. (2003). Vagal influence on working memory and attention. *International Journal of Psychophysiology, 48*(3), 263-274.

Hare, R. D. (1973). Orienting and defensive responses to visual stimuli. *Psychophysiology, 10*(5), 453-464.

Harris, C. L. (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development, 25*(2-3), 223-247.

Harris, C. L., Aycicegi, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics, 24*(04), 561-579.

Hayano, J., Sakakibara, Y., Yamada, A., Yamada, M., Mukai, S., Fujinami, T., . . . Takata, K. (1991). Accuracy of assessment of cardiac vagal tone by heart rate variability in normal subjects. *American Journal of Cardiology, 67*(2), 199-204.

Heck, R. (2016). Locating multivariate outliers. Retrieved from file:///Users/Jessica/Downloads/Locating%20Multivariate%20Outliers%20(1).pdf

Hilton, S. (1975). Ways of viewing the central nervous control of the circulation—old and new. *Brain Research, 87*(2), 213-219.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53-60.

Horwitz, E. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics, 21*, 112-126.

Horwitz, E., Horwitz, M. B., & Cope, J. A. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125-132.

Horwitz, E., & Young, D. J. (1991). *Language anxiety: From theory and research to classroom implications*. New York City: Pearson College Division.

Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching, 43*(02), 154. doi:10.1017/s026144480999036x

Howell, D. C. (2012). *Statistical methods for psychology*. Independence, KY: Cengage Learning.

Huang-Pollock, C. L., Carr, T. H., & Nigg, J. T. (2002). Development of selective attention: Perceptual load influences early versus late attentional selection in children and adults. *Developmental Psychology, 38*(3), 363-375.

Hugdahl, K. (1995). *Psychophysiology: The mind-body perspective*. Cambridge: Harvard University Press.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539-558.

Izard, C. E. (1989). The structure and functions of emotions: Implications for cognition, motivation, and personality. In I. S. Cohen (Ed.), *The G. Stanley Hall lecture series* (Vol. 9, pp. 39-73). Washington, DC: American Psychological Association.

Izard, C. E. (1997). Emotions and facial expressions: A perspective from Differential Emotions Theory. In J. A. Russell & J. M. Fernández-Dols (Eds.), *The psychology of facial expression* (Vol. 2, pp. 57-77). Cambridge: Cambridge University Press.

Jaén, M. M., & Basanta, C. P. (2009). Developing conversational competence through language awareness and multimodality: The use of DVDs. *ReCALL, 21*(03), 283-301.

Jönsson, P. (2007). Respiratory sinus arrhythmia as a function of state anxiety in healthy individuals. *International Journal of Psychophysiology, 63*(1), 48-54.

Kameenui, E. J., Carnine, D. W., & Freschi, R. (1982). Effects of text construction and instructional procedures for teaching word meanings on comprehension and recall. *Reading Research Quarterly, 17*(3), 367-388.

Karenbach, C. (2005). *Ledalab-a software package for the analysis of phasic electrodermal activity*. Retrieved from

Keogh, E., & Reidy, J. (2000). Exploring the factor structure of the Mood and Anxiety Symptom Questionnaire (MASQ). *Journal of personality assessment, 74*(1), 106-125. doi:10.1207/S15327752JPA740108

Khan, Z. A., & Zafar, S. (2010). The effects of anxiety on cognitive processing in English language learning. *English Language Teaching, 3*(2), 199-209.

Kihlstrom, J. F., Eich, E., Sandbrand, D., & Tobias, B. A. (2000). *Emotion and memory: Implications for self-report*. New York: Psychology Press.

Klabunde, R. (Producer). (2016). Determining heart rate from the electrocardiogram. *Cardiovascular Physiology Concepts*. Retrieved from http://www.cvphysiology.com/Arrhythmias/A020

Kleinmann, H. H. (1977). Avoidance behavior in adult second language acquisition. *Language Learning, 27*(1), 93-107.

Kleinsmith, L. J., & Kaplan, S. (1963). Paired-associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology, 65*(2), 190.

Kleinsmith, L. J., & Kaplan, S. (1964). Interaction of arousal and recall interval in nonsense syllable paired-associate learning. *Journal of Experimental Psychology, 67*(2), 124.

Kleinsmith, L. J., Kaplan, S., & Trate, R. D. (1963). The relationship of arousal to short-and long-term verbal recall. *Canadian Journal of Psychology, 17*(4), 393.

Klorman, R. (1974). Habituation of fear: Effects of intensity and stimulus order. *Psychophysiology, 11*(1), 15-26.

Klorman, R., Weissberg, R. P., & Wiesenfeld, A. R. (1977). Individual differences in fear and autonomic reactions to affective stimulation. *Psychophysiology, 14*(1), 45-51.

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals, 22*(6), 529-540.

Kok, B., & Fredrickson, B. (2010). Upward spirals of the heart: Autonomic flexibility, as indexed by vagal tone, reciprocally and prospectively predicts positive emotions and social connectedness. *Biological psychology, 85*(3), 432-436.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. Boston: Addison-Wesley Longman Ltd.

Krashen, S., Dulay, H., & Burt, M. (1982). *Language two*. New York: Oxford University Press.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological psychology, 84*(3), 394-421. doi:10.1016/j.biopsycho.2010.03.010

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. doi:10.1111/j.1469-8986.2007.00550.x

Lacey, B. C., & Lacey, J. I. (1974). Studies of heart rate and other bodily processes in sensorimotor behavior. In P. A. Obrist, A. H. Black, J. Brener, & L. V. DiCara (Eds.), *Cardiovascular psychophysiology: Current issues in response mechanisms, biofeedback and methodology* New Brunswick, NJ: AldineTransaction.

Lacey, J. I. (1959). *Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome.* Paper presented at the Research in Psychotherapy, Washington, DC.

Lacey, J. I., & Lacey, B. C. (1958). Verification and extension of the principle of autonomic response-stereotypy. *The American journal of psychology, 71*(1), 50-73.

Lacey, J. I., & Lacey, B. C. (1974). On heart rate responses and behavior: a reply to Elliott. *Journal of personality and social psychology, 30*(1), 1-18.

Lagere, M. (2015). Department of Defense strategic language list. Retrieved from http://www.altalang.com/wp-content/uploads/2015/06/FY15-Army-Strategic-Language-List-ASLL.pdf

Lane, R. D., Nadel, L., Allen, J. J., & Kaszniak, A. W. (2000). The study of emotion from the perspective of cognitive neuroscience. *Cognitive neuroscience of emotion*, 3-11.

Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology, 16*(6), 495-512.

Lang, P. J., & Bradley, M. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. In J. Coan & J. Allen (Eds.), *Handbook of emotion elicitation and assessment* (Vol. 29). Oxford: Oxford University Press.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological review, 97*(3), 377-395.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. *Attention and orienting: Sensory and motivational processes*, 97-135.

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. doi:DOI 10.1111/j.1469-8986.1993.tb03352.x

Lang, P. J., Simons, R. F., & Balaban, M. (2013). *Attention and orienting: Sensory and motivational processes*. New York: Psychology Press.

Larson-Hall, J. (2009). A guide to doing statistics in second language research rsing SPSS. New York: Routledge.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Palgrave Macmillan UK.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. *Second Language Vocabulary Acquisition*, 3034.

Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA yearbook, 5*(1), 223-250. doi:DOI 10.1075/eurosla.5.11lau

Lazarus, R. S., & Folkman, S. (1984). Coping and adaptation. In W. W. Gentry (Ed.), *The handbook of behavioral medicine* (pp. 282-325). New York: The Guilford Press.

Lazarus, R. S., Speisman, J. C., & Mordkoff, A. M. (1963). The relationship between autonomic indicators of psychological stress: Heart rate and skin conductance. *Psychosomatic Medicine, 25*(1), 19-30.

Lazarus, R. S., Speisman, J. C., Mordkoff, A. M., & Davison, L. A. (1962). A laboratory study of psychological stress produced by a motion-picture film. *Psychological Monographs, 76*(34), 1-35.

Lehrer, P. M., & Gevirtz, R. (2014). Heart rate variability biofeedback: how and why does it work? *Frontiers in Psychology, 5*, 756.

Leichtman, M. D., Ceci, S. J., & Ornstein, P. A. (1992). The influence of affect on memory: Mechanism and development. *The handbook of emotion and memory: Research and theory*, 181-199.

Leow, R. P. (1998). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) finegrained analysis of attention. *Applied Psycholinguistics, 19*(01), 133-159.

Levenson, R. W. (1992). Autonomic nervous-system differences among emotions. *Psychological science, 3*(1), 23-27. doi:DOI 10.1111/j.1467-9280.1992.tb00251.x

Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology, 27*(4), 363-384.

Levitt, E. E. (1980). *The psychology of anxiety*. Mahwah, NJ Lawrence Erlbaum Associates, Inc.

Lewis, M. (1993). *The lexical approach*. Hove, UK: Language Teaching Publications.

Libby, W. L., Jr., Lacey, B. C., & Lacey, J. I. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology, 10*(3), 270-294.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching. *Studies in Second Language Acquisition, 12*(04), 429-448.

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., . . . Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning, 63*(3), 530-566.

Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal, 92*(1), 71-86.

Loewen, S., & Gonulal, T. (2015). Factor analysis. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 182-212). New York: Routledge.

Long, M. H. (1983). Does second language instruction make a difference? A review of research. *Tesol Quarterly, 17*(3), 359-382.

Lotto, L., & de Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning, 48*(1), 31-69. doi:Doi 10.1111/1467-9922.00032

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

MacIntyre, P. (1994). Variables underlying willingness to communicate: A causal analysis. *Communication Research Reports, 11*(2), 135-142.

MacIntyre, P. (1995). How does anxiety affect second language learning? A reply to Sparks and Ganschow. *The Modern Language Journal, 79*(1), 90-99.

MacIntyre, P. (2002). Motivation, anxiety and emotion in second language acquisition. *Individual differences and instructed language learning, 2*, 45-68.

MacIntyre, P. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. . *The Modern Language Journal, 91*(4), 564-576.

MacIntyre, P. (2012). The idiodynamic method: A closer look at the dynamics of communication traits. *Communication Research Reports, 29*(4), 361-367. doi:10.1080/08824096.2012.723274

MacIntyre, P., & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of language and social psychology, 15*(1), 3-26.

MacIntyre, P., & Gardner, R. C. (1989). Anxiety and second-language learning: Toward a theoretical clarification*. *Language Learning, 39*(2), 251-275.

MacIntyre, P., & Gardner, R. C. (1991). Methods and results in the study of anxiety and language learning: A review of the literature. *Language Learning, 41*(1), 85-117.

MacIntyre, P., & Gardner, R. C. (1994a). The effects of induced anxiety on three stages of cognitive processing in computerized vocabulary learning. *Studies in Second Language Acquisition, 16*(1), 1-17.

MacIntyre, P., & Gardner, R. C. (1994b). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*(2), 283-305.

Macintyre, P., & Legatto, J. J. (2010). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied linguistics, 32*(2), 149-171. doi:10.1093/applin/amq037

Malmstrom, E. J., Opton, E., Jr., & Lazarus, R. S. (1965). Heart rate measurement and the correlation of indices of arousal. *Psychosomatic Medicine, 27*(6), 546-556.

Marcos-Llinás, M., & Garau, M. J. (2009). Effects of language anxiety on three proficiency-level courses of Spanish as a foreign language. *Foreign Language Annals, 42*(1), 94-111.

Marker, C. (Producer). (2017). Heart rate variability: The new science of recovery. Retrieved from https://breakingmuscle.com/learn/heart-rate-variability-the-new-science-of-recovery

Marschark, M., Richman, C. L., Yuille, J. C., & Hunt, R. R. (1987). The role of imagery in memory: On shared and distinctive information. *Psychological bulletin, 102*(1), 28-41.

Martin, H. (Producer). (2012). Engaging the intelligence of the heart. *TEDx Santa Cruz*. Retrieved from https://www.youtube.com/watch?v=A9kQBAH1nK4

Mateo, J., & Laguna, P. (2003). Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Transactions on Biomedical Engineering, 50*(3), 334-343. doi:10.1109/Tbme.2003.808831

Mates, A., & Joaquin, A. (2013). Affect and the brain. In J. Herschensohn & M. Young–Scholten (Eds.), *Cambridge handbook of second language acquisition* (pp. 417-435). Cambridge Cambridge University Press.

Mathieu, J. E., Tannenbaum, S. I., & Salas, E. (1992). Influences of individual and situational characteristics on measures of training effectiveness. *Academy of management journal, 35*(4), 828-847. doi:Doi 10.2307/256317

MATLAB User's Guide. (1998). (Vol. 5). Natick, MA: The mathworks, Inc.

Matlin, M. W., & Stang, D. J. (1978). *The Pollyanna principle: Selectivity in language, memory, and thought*. Rochester, VT: Schenkman Pub. Co.

Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System, 32*(1), 21-36.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion, 23*(2), 209-237. doi:10.1080/02699930802204677

McLean, P. (1969). Induced arousal and time of recall as determinants of paired-associate recall *British journal of Psychology, 60*(1), 57-62.

Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching, 13*(3-4), 221-246.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 28*(6), 487-495.

Michel, L. (Producer). (2016). Heart rate variability science meets heart and mindful practices. *Talks at Google*. Retrieved from https://www.youtube.com/watch?v=GNYSnDQ8yAc

Miller, Z. (2016). *Playing with fire: Effects of hot cognition on working memory and L2 vocabulary acquisition*. Second Language Studies Program Qualifying Research Paper. Michigan State University.

Miller, Z., Fox, J., Moser, J., & Godfroid, A. (2017). *Playing with fire: Effects of hot cognition on working memory and L2 vocabulary acquisition*.

Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals, 39*(2), 276-295.

Moser, J. (2016). [Missing data of physiology measures].

Mueller, J. (1992). Anxiety and performance. In D. Jones & A. Smith (Eds.), *Handbook of human performance: State and trait* (Vol. 3, pp. 127-160). Ann Arbor, MI: Academic PRess.

Mundy-Castle, A., & McKiever, B. (1953). The psychophysiological significance of the galvanic skin response. *Journal of Experimental Psychology, 46*(1), 15.

Nation, I. S. P. (1993). Vocabulary size, growth, and use. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 115-134). Amsterdam: John Benjamins Publishing.

Nation, I. S. P. (2013). *Teaching & learning vocabulary*. Boston: Heinle Cengage Learning.

Nation, P., & Coady, J. (1988). Vocabulary and reading. *Vocabulary and Language Teaching, 97*(1), 110.

Onwuegbuzie, A. J., Bailey, P., & Daley, C. E. (2000). Cognitive, affective, personality, and demographic predictors of foreign-language achievement. *The Journal of Educational Research, 94*(1), 3-15. doi:10.1080/00220670009598738

Osborne, J. W. (2008). *Best practices in quantitative methods*. Thousand Oaks, CA: Sage.

Oveis, C., Cohen, A. B., Gruber, J., Shiota, M. N., Haidt, J., & Keltner, D. (2009). Resting respiratory sinus arrhythmia is associated with tonic positive emotionality. *Emotion, 9*(2), 265.

Oxford, R., & Crookall, D. (1989). Research on language learning strategies: Methods, findings, and instructional issues. *The Modern Language Journal, 73*(4), 404-419.

Oxford, R. L. (2014). What we can learn about strategies, language learning, and life from two extreme cases: The role of well-being theory. *Studies in Second Language Learning and Teaching, 4*(4), 593. doi:10.14746/ssllt.2014.4.4.2

Pagani, M., Furlan, R., Pizzinelli, P., Crivellaro, W., Cerutti, S., & Malliani, A. (1989). Spectral analysis of RR and arterial pressure variabilities to assess sympatho-vagal interaction during mental stress in humans. *Journal of Hypertension, 7*(6), S14-15.

Pagani, M., Mazzuero, G., Ferrari, A., Liberati, D., Cerutti, S., Vaitl, D., . . . Malliani, A. (1991). Sympathovagal interaction during mental stress: A study using spectral analysis of heart rate variability in healthy control subjects and patients with a prior myocardial infarction. *Circulation, 83*(4 Suppl), II43-51.

Paivio, A., & Desrochers, A. (1981). Mnemonic techniques in second-language learning. *Journal of Educational Psychology, 73*(6), 780-795.

Palomba, D., Sarlo, M., Angrilli, A., Mini, A., & Stegagno, L. (2000). Cardiac responses associated with affective processing of unpleasant film stimuli. *International Jounral of Psychophysiology, 36*(1), 45-57.

Park, G., Van Bavel, J. J., Vasey, M. W., & Thayer, J. F. (2012). Cardiac vagal tone predicts inhibited attention to fearful faces. *Emotion, 12*(6), 1292-1302. doi:10.1037/a0028528

Park, G., Vasey, M., Van Bavel, J., & Thayer, J. (2013). Cardiac vagal tone is correlated with selective attention to neutral distractors under load. *Psychophysiology, 50*(4), 398-406.

Park, G., Vasey, M. W., Van Bavel, J. J., & Thayer, J. F. (2014). When tonic cardiac vagal tone predicts changes in phasic vagal tone: the role of fear and perceptual load. *Psychophysiology, 51*(5), 419-426. doi:10.1111/psyp.12186

Pauls, C. A., & Stemmler, G. (2003). Repressive and defensive coping during fear and anger. *Emotion, 3*(3), 284-302. doi:10.1037/1528-3542.3.3.284

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition, 11*(02), 147-164.

Peltola, M. A. (2007). Role of editing of R–R intervals in the analysis of heart rate variability. *Heart Rate Variability: Clinical Applications and Interaction between HRV and Heart Rate*, 29.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage.

Pimsleur, P. (1966). *Pimsleur Language Aptitude Battery (form S)*: Harcourt, Brace and world, Incorporated.

Plonsky, L. (2013). Study quality in SLA. *Studies in Second Language Acquisition, 35*(04), 655-687.

Plonsky, L. (2014). Study quality in quantitative L2 research (1990–2010): A methodological synthesis and call for reform. *The Modern Language Journal, 98*(1), 450-470.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Theories of Emotion, 1*, 3-31.

Poels, K., & Dewitte, S. (2006). How to capture the heart? Reviewing 20 years of emotion measurement in advertising. *Journal of Advertising Research, 46*(1), 18-37.

Porges, S. W. (1986). Respiratory sinus arrhythmia: Physiological basis, quantitative methods, and clinical implications *Cardiorespiratory and cardiosomatic psychophysiology* (pp. 101-115). New York: Springer.

Porges, S. W. (1995). Cardiac vagal tone: A physiological index of stress. *Neuroscience & Biobehavioral Reviews, 19*(2), 225-233.

Porges, S. W., Doussard-Roosevelt, J. A., & Maiti, A. K. (1994). Vagal tone and the physiological regulation of emotion. *Monographs of the Society for Research in Child Development, 59*(2-3), 167-186.

Reidy, J., & Keogh, E. (1997). Testing the discriminant and convergent validity of the Mood and Anxiety Symptoms Questionnaire using a British sample. *Personality and individual Differences, 23*(2), 337-344.

Reis, H. T., & Judd, C. M. (2000). *Handbook of research methods in social and personality psychology*. Cambridge: Cambridge University Press.

Revelle, W., & Loftus, D. A. (1992). The implications of arousal effects for the study of affect and memory. *The handbook of emotion and memory: Research and theory*, 113-149.

Robinson, M. D., & Clore, G. L. (2002). Episodic and semantic knowledge in emotional self-report: evidence for two judgment processes. *Journal of personality and social psychology, 83*(1), 198-215.

Robinson, M. D., & Sedikides, C. (2009). Traits and the self: Toward an integration. In P. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality psychology* (pp. 457-472). Cambridge: Cambridge University Press.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*(2), 283-331.

Robinson, P. (2002). Learning conditions, aptitude complexes, and SLA. *Individual differences and instructed language learning*, 113-133.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. (Version Version 0.5–12 (BETA)). Ghent, Belgium: Ghent University.

Santosa, D. (Producer). (2014). Ambon island map. *Travel Donk*. Retrieved from http://www.traveldonk.com/2013/02/a-hidden-beach-in-eastern-ambon-naku.html

Sarason, I. G. (1986). Test anxiety, worry, and cognitive interference. In R. Schwarzer (Ed.), *Self-related cognitions in anxiety and motivation* (1 ed., pp. 19-33). New York: Psychology Press.

Schein, E. H. (2002). The anxiety of learning. Interview by Diane L. Coutu. *Harvard Business Review, 80*(3), 100-106, 134.

Scherer, K. R., Scherer, K. R., & Ekman, P. (1984). On the nature and function of emotion: A component process approach. *Approaches to Emotion, 2293*, 317.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of personality and social psychology, 85*(3), 440.

Schmidt, R. (1992). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*(1), 206-226.

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329-363.

Schmitt, N. (2010a). *Key issues in teaching and learning vocabulary*. Bristol: Multilingual Matters.

Schmitt, N. (2010b). *Researching vocabulary: A vocabulary research manual*. New York: Springer.

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 1-39.

Schneiderman, N., & McCabe, P. M. (1989). Psychophysiologic strategies in laboratory research. In N. Schneiderman, S. Weiss, & P. Kaufmann (Eds.), *Handbook of research methods in cardiovascular behavioral medicine* (pp. 349-364). New York: Springer.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*(6), 323-338.

Schumann, J. H. (1997). The neurobiology of affect in language. *Language Learning, 48*.

Schumann, J. H. (1999). A neurobiological perspective on affect and methodology in second language learning. *Affect in language learning*, 28-42.

Schumann, J. H., Crowell, S. E., Jones, N. E., Lee, N., & Schuchert, S. A. (2014). *The neurobiology of learning: Perspectives from second language acquisition*. Abingdon-on-Thames: Routledge.

Seidenberg, M., Haltiner, A., Taylor, M. A., Hermann, B. B., & Wyler, A. (1994). Development and validation of a Multiple Ability Self-Report Questionnaire. *Journal of Clinical and Experimental Neuropsychology, 16*(1), 93-104. doi:10.1080/01688639408402620

Sequeira, H., Hot, P., Silvert, L., & Delplanque, S. (2009). Electrical autonomic correlates of emotion. *International Journal of Psychophysiology, 71*(1), 50-56. doi:10.1016/j.ijpsycho.2008.07.009

Shaver, P. R., Morgan, H. J., & Wu, S. (1996). Is love a "basic" emotion? *Personal Relationships, 3*(1), 81-96.

Simard, D., & Wong, W. (2001). Alertness, orientation, and detection. *Studies in Second Language Acquisition, 23*(01), 103-124.

Sinha, R., Lovallo, W. R., & Parsons, O. A. (1992). Cardiovascular differentiation of emotions. *Psychosomatic Medicine, 54*(4), 422-435.

Skin Conductance Explained. (2017). *Psychlab.* Retrieved from http://www.psychlab.com/SC_explained.html

Smith, C. D., Sawyer, C. R., & Behnke, R. R. (2005). Physical symptoms of discomfort associated with worry about giving a public speech. *Communication Reports, 18*(1-2), 31-41. doi:10.1080/08934210500084206

Sokolov, E. N. (1963). Higher nervous functions; the orienting reflex. *Annual review of physiology, 25*(1), 545-580. doi:10.1146/annurev.ph.25.030163.002553

Sparks, R., & Ganschow, L. (2007). Is the Foreign Language Classroom Anxiety Scale measuring anxiety or language skills? *Foreign Language Annals, 40*(2), 260-287.

Spielberger, C. D. (2010). *State-Trait anxiety inventory*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Spielberger, C. D., & Spielberger, C. D. (1966). Theory and research on anxiety. *Anxiety and behavior, 1*, 3-20.

Spielmann, G., & Radnofsky, M. L. (2001). Learning language under tension: New directions from a qualitative study. *The Modern Language Journal, 85*(2), 259-278.

Srivastava, S., Guglielmo, S., & Beer, J. S. (2010). Perceiving others' personalities: Examining the dimensionality, assumed similarity to the self, and stability of perceiver effects. *Journal of personality and social psychology, 98*(3), 520-534. doi:10.1037/a0017057

Steinberg, F., & Horwitz, E. (1986). The effect of induced anxiety on the denotative and interpretive content of second language speech. *Tesol Quarterly, 20*(1), 131-136.

Stemmler, G. (1989). The autonomic differentiation of emotions revisited: Convergent and discriminant validation. *Psychophysiology, 26*(6), 617-632.

Sternbach, R. A. (1962). Assessing differential autonomic patterns in emotions. *Journal of psychosomatic research, 6*(2), 87-91.

Stone, A. A., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S., & Cain, V. S. (1999). *The science of self-report: Implications for research and practice*. New York: Psychology Press.

Suen, H., & Ary, D. (1989). Analyzing quantitative observation data. Hillsdale, NJ: Lawrence Erlbaum Associates.

Surwillo, W. W. (1990). *Psychophysiology for clinical psychologists*. New York: Ablex Publishing.

Suspension. (2017). Suspension bridge. *Parabolas in Suspension Bridges!* Retrieved from http://www.carondelet.pvt.k12.ca.us/Family/Math/03210/page4.htm

Swain, M. (2011). The inseparability of cognition and emotion in second language learning. *Language Teaching, 46*(02), 195-207. doi:10.1017/s0261444811000486

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., . . . Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*(2), 247-250.

Tarvainen, M., & Niskanen, J. (2012). Kubios HRV. Finland, Department of Applied Physics, University of Eastern Finland.

Taylor, T. (2016). Nervous system. *Inner Body.* Retrieved from http://www.innerbody.com/image/nervov.html#full-description

Team, R. C. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Thayer, J. F., Friedman, B. H., & Borkovec, T. D. (1996). Autonomic characteristics of generalized anxiety disorder and worry. *Biological Psychiatry, 39*(4), 255-266. doi:10.1016/0006-3223(95)00136-0

Thayer, J. F., Hansen, A. L., & Johnsen, B. H. (2010). The non-invasive assessment of autonomic influences on the heart using impedance cardiography and heart rate variability *Handbook of behavioral medicine* (pp. 723-740). New York: Springer.

Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. *Annals of Behavioral Medicine, 37*(2), 141-153. doi:10.1007/s12160-009-9101-z

Thayer, J. F., & Lane, R. D. (2009). Claude Bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience & Biobehavioral Reviews, 33*(2), 81-88.

Tobias, S. (1986). Anxiety and cognitive processing of instruction. *Self-related Cognition in Anxiety and Motivation*, 35-54.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(02), 183-203.

Tracy, J. L., & Robins, R. W. (2004). Show your pride evidence for a discrete emotion expression. *Psychological science, 15*(3), 194-197.

Tran, T. Q., Boring, R. L., Dudenhoeffer, D. D., Hallbert, B. P., Keller, M. D., & Anderson, T. M. (2007). *Advantages and disadvantages of physiological assessment for next generation control room design.* Paper presented at the Human Factors and Power Plants and HPRCT 13th Annual Meeting.

Tremayne, P., & Barry, R. J. (2001). Elite pistol shooters: Physiological patterning of best vs. worst shots. *International Journal of Psychophysiology, 41*(1), 19-29.

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

Utara, B. I. M. M. (2016). Completed bridge.

Vazire, S., & Carlson, E. N. (2011). Others sometimes know us better than we know ourselves. *Current Directions in Psychological Science, 20*(2), 104-108.

Verhees, M. W., Chwilla, D. J., Tromp, J., & Vissers, C. T. (2015). Contributions of emotional state and attention to the processing of syntactic agreement errors: Evidence from P600. *Frontiers in Psychology, 6*, 388. doi:10.3389/fpsyg.2015.00388

Vianna, E. P., & Tranel, D. (2006). Gastric myoelectrical activity as an index of emotional arousal. *International Journal of Psychophysiology, 61*(1), 70-76.

Viinikainen, M., Glerean, E., Jääskeläinen, I. P., Kettunen, J., Sams, M., & Nummenmaa, L. (2012). Nonlinear neural representation of emotional feelings elicited by dynamic naturalistic stimulation. *Open Journal of Neuroscience, 2*(4), 1-17.

Volokhov, R. N., & Demaree, H. A. (2010). Spontaneous emotion regulation to positive and negative stimuli. *Brain and Cognition, 73*(1), 1-6. doi:10.1016/j.bandc.2009.10.015

Warr, P., & Downing, J. (2000). Learning strategies, learning anxiety and knowledge acquisition. *British journal of Psychology, 91 (Pt 3)*(3), 311-333.

Watson, D., & Clark, L. A. (1991). The Mood and Anxiety Symptom Questionnaire. Iowa City: University of Iowa.

Watson, J. (Producer). (2015). Heart rate variability: Principles and science and practical measuring devices Retrieved from http://www.anti-agingfirewalls.com/__oneclick_uploads/2015/07/Med3.jpg

Watson, P. J., & Gatchel, R. J. (1979). Autonomic measures of advertising. *Journal of Advertising Research, 19*(3), 15-26.

West, S. G., & Brown, T. J. (1975). Physical attractiveness, the severity of the emergency and helping: A field experiment and interpersonal simulation. *Journal of Experimental Social Psychology, 11*(6), 531-538.

Wilkins, D. A. (1972). *Linguistics in language teaching*. London: Edward Arnold.

Williams, R. (2016). Outliers. Retrieved from https://www3.nd.edu/~rwilliam/stats2/l24.pdf

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*(1), 65-86.

Winton, W. M., Putnam, L. E., & Krauss, R. M. (1984). Facial and autonomic manifestations of the dimensional structure of emotion. *Journal of Experimental Social Psychology, 20*(3), 195-216.

Witliet, C., & Vrana, S. (1995). Psychophysiological responses as indices of affective dimensions. *Psychophysiology, 32*(5), 436-443.

Witt, P. L., Brown, K. C., Roberts, J. B., Weisel, J., Sawyer, C. R., & Behnke, R. R. (2006). Somatic anxiety patterns before, during, and after giving a public speech. *Southern Communication Journal, 71*(1), 87-100. doi:10.1080/10417940500503555

Wolfe, P. (2006). The role of meaning and emotion in learning. *New directions for adult and continuing education, 2006*(110), 35-41.

Young, D. J. (1994). New directions in language anxiety research. In C. A. Klee (Ed.), *Faces in a crowd: The individual learner in multisection courses*. Boston: Heinle & Heinle Publishers.

Zeidner, M. (1998). *Test anxiety: The state of the art*. New York: Springer Science & Business Media.