Fear Conditioning as an Intermediate Phenotype: An RDoC Inspired Methodological Analysis

Michael W. Lewis

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Psychology

Russell T. Jones

Bruce H. Friedman

Richard A. Winett

4/20/2018

Blacksburg, VA

Keywords: Fear Learning, Threat Conditioning, Latent Growth Mixture Modeling, Latent Class Growth Analysis

Fear Conditioning as an Intermediate Phenotype: An RDoC Inspired Methodological Analysis

Michael W. Lewis

ABSTRACT

Due to difficulties in elucidating neurobiological aspects of psychological disorders, the National Institute of Mental Health (NIMH) created the Research Domain Criteria (RDoC), which encourages novel conceptualizations of the relationship between neurobiological circuitry and clinical difficulties. This approach is markedly different from the Diagnostic and Statistical Manual of Mental Disorders (DSM) based approach that has dominated clinical research to date. Thus, RDoC necessitates exploration of novel experimental and statistical approaches. Fear learning paradigms represent a promising methodology for elucidating connections between acute threat ("fear") circuitry and fear-related clinical difficulties. However, traditional analytical approaches rely on central tendency statistics, which are tethered to a priori categories and assume homogeneity within groups. Growth Mixture Modeling (GMM) methods such as Latent Class Growth Analysis (LCGA) may be uniquely suited for examining fear learning phenotypes. However, just three extant studies have applied GMM to fear learning and only one did so in a human population. Thus, the degree to which classes identified in known studies represent characteristics of the general population and to which GMM methodology is applicable across populations and paradigms is unclear. This preliminary study applied LCGA to a fear learning lab study in an attempt to identify heterogeneity in fear learning patterns based on a posteriori classification. The findings of this investigation may inform efforts to move toward a trans-diagnostic conceptualization of fear learning. Consistent with the goals laid out in RDoC, explication of fear learning phenotypes may eventually provide critical information needed to spur innovation in psychotherapeutic and psychopharmacological treatment.

Fear Conditioning as an Intermediate Phenotype: An RDoC Inspired Methodological Analysis

Michael W. Lewis

ABSTRACT

To date, most clinical psychology research has been based on the Diagnostic and Statistical Manual of Mental Disorders (DSM), which is a catalog of mental health disorders that was originally designed to facilitate communication among clinicians. Many experts contend that this approach has hampered progress in the field of biological clinical psychology research. Thus, the National Institute of Mental Health (NIMH) created a new template for biological clinical psychology research called the Research Domain Criteria (RDoC). Since RDoC calls for a complete overhaul in the conceptualization of clinical dysfunction, this approach requires statistical and experimental innovation. One traditional experimental approach that may be helpful in understanding the RDoC topic of acute threat ("fear") is called Pavlovian Fear Learning (PFL). However, traditional PFL studies have utilized statistical methods that are based on comparing group averages and require researchers to determine groups of interest based on theory before the study begins. This is problematic because RDoC calls for research that begins with evidence rather than theory. Growth Mixture Modeling (GMM) is a statistical methodology that may allow researchers to analyze fear learning data without having to begin with theoretically determined categories such as DSM disorders. However, little research has tested how well this approach would work. This study is just the second to apply a GMM approach to a human PFL study. The findings from this investigation may inform efforts to develop a statistical technique that is well suited for RDoCian research and may also spur innovation in psychotherapeutic and psychopharmacological treatment.

ACKNOWLEDGMENTS

First, thank you to my advisor, Russell T. Jones, for his guidance and mentorship during my first three years of graduate school, including this thesis process. I cannot tell you how grateful I am that you believed in me and supported my ambition to seek a Ph.D. in psychology and take on a challenging project such as this.

Second, thank you to Bruce H. Friedman, for adopting me into the Mind Body lab as an unofficial honorary member and providing additional support and guidance in my transition into Biological Psychology research throughout the past two years.

Third, thank you to Richard Winnett for providing additional input and mentorship both professionally and personally.

Finally, thank you to all of the graduate and undergraduate students who helped in the data collection process and made this project possible.

Table of Contents

List of Tables and Figures

1.0 Introduction

Attempts to elucidate the relationship between dysfunctional neurobiological processes and symptomatic thoughts, feelings, and behaviors which characterize psychological disorders have yielded inconsistent results, especially in disorders with a wide array of possible symptom combinations such as generalized anxiety disorder (GAD), panic disorder (PD), and posttraumatic stress disorder (PTSD) (Kozak & Cuthbert, 2016; Galatzer-Levy & Bryant, 2013). The NIMH RDoC project encourages researchers to examine neurobiological processes, thoughts, emotions, health, and behavior using a bottom-up approach across the entire spectrum of wellbeing, rather than focusing on diagnostic groups (Kozak & Cuthbert, 2016).

Since RDoC discourages researchers from categorizing subjects based on DSM disorders and promotes new conceptualizations that are built upon neurobiological evidence, researchers should employ experimental methodologies that index well understood neurobiological circuitry and are likely to be related to clinical problems across disorders (Kozak & Cuthbert, 2016). One such methodology is Pavlovian Fear Learning (PFL), which utilizes fear learning paradigms, and indexes processes through which the aversive properties of a naturally harmful or unpleasant stimulus come to be paired with the sensory properties of an otherwise neutral or pleasant stimulus as well as the process through which these associations are inhibited or extinguished (Davis, 2006). PFL is theorized to mediate a wide array of symptoms and to modulate others, but empirical findings of PFL in DSM disorders have been inconsistent, which may suggest a need for more precise analytical techniques (Galatzer-Levy & Bryant, 2013; Zoellner et al., 2014).

In order to apply experimental indices of neurobiological function to a paradigm which seeks to revolutionize biological clinical psychology research, researchers must test statistical

techniques which may support an entirely new way of conceptualizing clinical dysfunction. Recent studies suggest that Latent Growth Mixture Modeling (LGMM) can be used to identify heterogeneity in PFL that is missed in studies which utilize more traditional statistical approaches (Galatzer-Levy et al., 2013 b.; Galatzer-Levy et al., 2017). However, additional studies are needed to further validate this approach and refine hypotheses regarding fear learning phenotypes. Thus, this study evaluated latent profiles of fear extinction, fear extinction recall, and safety signal learning trajectories in a human sample by analyzing FPS during a fear learning paradigm that utilizes methodological approaches that are yet to be tested in a GMM study.

1.1 Research Domain Criteria (RDoC)

The National Institute of Mental Health Research Domain Criteria (NIMH RDoC) is a promising new clinical research paradigm that may allow researchers to study constructs that are more readily linked with biological systems than are DSM disorders. The RDoC framework is empirically driven, agnostic to DSM categories, dimensional in measurement, and oriented towards biological elaboration of narrowly defined clinical problems using intermediate constructs and multiple units of analysis (Insel et al., 2010; Kozak & Cuthbert, 2016). RDoC constructs require: persuasive empirical evidence for their validity, tangible evidence of a neural circuit or system which implements the construct, and a theoretical link to a relevant clinical phenomenon (Kozak & Cuthbert, 2016). Thus, this system takes a bottom-up approach to investigating clinically relevant patterns of thoughts, emotions, and behaviors.

Research investigating the biological aspects of fear-related DSM disorders has yielded inconsistent results. So far, neurobiological and genetics research attempting to map neural circuitry onto DSM disorders have found inconsistent results (Kozak & Cuthbert, 2016). The

search for biological mechanisms of PTSD has been especially vexing as findings have varied and often conflicted (Sherin & Nemeroff, 2011). Evidence suggests that attempting to treat individuals with PTSD as a monolithic group is likely to impede progress in explicating biological aspects of the disorder (Lanius, Bluhm, Lanius, & Pain, 2006). The heterogeneity of neurobiological patterns found in PTSD may be greater than commonly stated and a multitude of biological and self-report findings are inconsistent with common conceptualizations of PTSD (for a review, see Zoladz & Diamond, 2013). In an effort to understand and address the lack of progress toward consensus regarding biological aspects of clinical phenomena, many have critiqued the constructs contained within the DSM and have turned to RDoC as a promising alternative.

The RDoC framework offers potential solutions to the DSM's shortcomings. Common criticisms of the DSM constructs include: theoretically defined constructs are prematurely reified, there is a lack of clear boundaries between constructs, constructs are heterogeneous, and dimensional phenomena are defined categorically. Due to RDoC's bottom-up approach to explicating narrowly defined clinical phenomenon using biological evidence, this approach helps address many of the difficulties in investigating biological components of DSM constructs such as PTSD and other fear disorders (Kozak & Cuthbert, 2016). The DSM construct PTSD serves as an informative case study for common criticisms of the DSM approach as well as the potential advantages of the RDoC approach. Thus, in this section, each criticism will be briefly reviewed, applied to PTSD, and contrasted with the RDoC construct of Acute Threat ("Fear").

DSM constructs are often prematurely reified while RDoC constructs are formulated using a bottom-up approach. For example, PTSD was introduced in DSM-III in 1987, one year

prior to the first epidemiologic study of psychological trauma, as a response to political pressure on the mental health field to recognize the mental health impacts of the Vietnam War and Concentration Camps (Helzer, Robins, & McEvoy, 1987; Centers for Disease Control Vietnam Experience Study Group, 1988). Rather than seeking evidence to explicate a theoretically reified disorder, RDoC constructs are formulated only after integrating persuasive biological evidence from multiple labs, which must explicate an underlying neural circuit or system. For example, the RDoC construct Acute Threat ("Fear") is inspired by a wealth of empirical research linking threat responses to psychological and behavioral difficulties as well as explicating the neuroscience of how these responses are implemented through specific neural cirucuits (Kozak & Cuthbert, 2016).

DSM disorders may not be clearly delineated from each other while RDoC was intentionally designed to maximize divergent validity. For example, DSM-5 added a new cluster of negative alterations in cognition and mood to PTSD, though this cluster may represent symptoms that are core to the construct of MDD and auxiliary to the concept of PTSD (Zoellner, Pruitt, Farach, & Jun, 2014). More broadly, anxiety disorders generally do not seem to represent distinct and restrictive groups (Lang, McTeague, & Bradley, 2016). In comparison, "RDoC workshops were organized around each domain so that overlaps and distinctions among constructs would be considered" (Kozak & Cuthbert, 2016, p. 289). Thus, the RDoC construct Acute Threat ("Fear") is appropriately differentiated from Loss and other constructs nested within the Negative Valence Systems domain.

DSM disorders are often highly heterogeneous while one goal of RDoC is to connect cross-validated constructs to narrowly defined clinical problems (Kozak & Cuthbert, 2016).

There are 636,120 unique combinations of PTSD. Thus, the construct of PTSD is defined in terms which allow for a nearly limitless array of motivational systems to contribute to the construct (Galatzer-Levy & Bryant, 2013; Norrholm et al., 2015). Conversely, RDoC is designed to encourage the isolated study of restricted clinical problems, rather than categorically grouping them (Kozak & Cuthbert, 2016). For example, Acute Threat ("Fear") is defined by "activation of the brain's defensive motivational system to promote behaviors that protect the organism from perceived danger" (NIMH).

DSM disorders are defined categorically based on counts and combinations of symptoms that are defined using binary cutoffs, while RDoC encourages dimensional analysis. DSM literature examining the effects of symptom severity is relatively scarce, though existing evidence suggests that some neurobiological variables are more readily linked with severity of particular symptoms or clusters than with DSM category (Norrholm et al., 2015; Norrholm et al., 2011). While much of the nascent RDoC literature investigates diagnostic categories, RDoC encourages investigation of "narrowly defined impairments of psychiatric clinical importance … [that is,] individual symptoms or very homogeneous symptom sets," in order to "free investigators from nosological categories." (Kozak & Cuthbert, 2016 p. 288, as cited by Patrick, 2016). For example, a dimension of Acute Threat ("Fear") could encompass excessive fear, flexibly adaptable fear responses, or pathological fearlessness. Thus, this single motivational system may contribute to phobic, healthy, or psychopathic behavioral phenotypes (Kozak & Cuthbert, 2016).

1.2 Fear Learning Paradigms and Fear Potentiated Startle

Fear learning paradigms are laboratory models commonly employed in studies of animals and humans to investigate patterns of conditioned fear responses, which are acquired by pairing an unconditioned aversive stimulus (UCS) repeatedly with a conditioned stimulus (CS). These paradigms can be used to investigate multiple aspects of fear learning, including fear acquisition (i.e. the acquisition of aversive responding to the CS itself), fear extinction (i.e. the ability to inhibit fear in the presence of a CS when it is no longer paired with the UCS), and fear extinction recall (i.e. the ability to inhibit fear in the presence of a CS by recalling previous extinction learning after time has passed). In spite of conceptual similarities and neurobiological overlap, these processes differ in meaningful ways both conceptually and neurologically. Thus, each must be experimentally isolated in order to examine the effects of each of these processes (Wendt, Neubert, Koenig, Thayer, & Hamm, 2015).

Fear acquisition requires new learning in which one's perception of an initially unaversive stimulus changes in response to repeatedly encountering this stimulus in tandem with an unconditioned stimulus (UCS) that is naturally aversive. Over time, one is conditioned to perceive the initially neutral stimulus as threatening, which is why it is labeled as a CS+ (i.e. Conditioned Stimulus with aversive conditioning). Fear extinction requires new learning in which the aversive UCS is no longer associated with the CS. This is tested in paradigms that include an acquisition phase (i.e. phase in which the UCS and CS are repeatedly paired) followed by an extinction phase (i.e. phase in which the CS+ is repeatedly presented without the UCS). This results in the CS+ acquiring both excitatory (during acquisition) and inhibitory (during extinction) associations with the UCS. Recall is tested by repeating the extinction phase (i.e. repeated presentations of CS+ without the UCS) of a paradigm after an experimentally

significant amount of time (e.g. one day) has passed. Some paradigms also include a CS that is never paired with an aversive UCS, which becomes associated with a lack of threat and acts as a safety signal; this is labeled as a CS- (i.e. Conditioned Stimulus without aversive conditioning) (Wendt et al., 2015).

Fear Potentiated Startle (FPS) is a common measure of fearful reactivity in fear learning paradigms and is defined as the increase in the frequency or magnitude of the reflexive startle response to an acoustic startle probe in the presence of a conditioned stimulus (Norrholm et al., 2015). Though FPS measures reflexive eye blink activity in orbicularis oculi (eye muscles) in response to a startle probe rather than brain activity, numerous studies in animal models and functional neuroimaging studies in humans validate it as a reliable indirect index of neural circuitry (Davis, 2006). NIMH lists fear learning paradigms as one of six suggested experimental tests for investigating Acute Threat ("Fear"). Additionally, they state that emotion-modulated startle paradigms may be interpreted as measures of neural circuits if they are validated by animal models or functional neuroimaging (nimh.nih.gov). Reflexive eye blink, which is commonly measured using facial electromyography (EMG), is the most easily measured and reliable behavioral manifestation of the startle response in humans (Grillon & Davis, 1997). Thus, fear learning paradigms measuring FPS utilizing facial EMG provide a suitable neurobiological index for RDoCian analysis.

1.3 Fear Learning Neurobiology

The neurobiological underpinnings of fear conditioning have been explicated through numerous rodent studies as well as functional neuroimaging studies in humans. Though there are overlaps, fear acquisition learning, extinction learning, and recall of extinction learning do not

rely upon identical neural systems. Thus, in order to understand each, experiments must either focus on just one or experimentally separate the effects of each within a single paradigm.

Fear conditioning results from experiences of a UCS paired with a CS (CS+). When presented with an unconditioned aversive stimulus (US), thalamic neurons transmit sensory information regarding the US and concurrent stimuli to the lateral amygdala (LA). This information is then processed and the US and CS are linked by the Basolateral Amygdala (BLA) complex, which includes the lateral, basal, and accessory-basal nuclei of the amygdala (Christianson, 2012). Fearful behaviors, such as increased eye-blinks, are enacted through excitatory pathways projecting from the thalamus and perihinal cortext to the basal amygdala (BA) and then to the central (CeA) and medial (MeA) nuclei of the amygdala. Then, the CeA and MeA project indirectly to the pontis caudalis (PnC), which is a specific nucleus of the brain stem and is in the acoustic startle pathway. From there, messages are transmitted to motor neurons and on to muscles, including facial muscles that are commonly used to capture startle using eyeblink. Rodent studies show that behavioral manifestations of fear learning such as FPS are partially mediated by glutamate acting on N-methyl-D-aspartate (NMDA) receptors in the PnC as well as intracellular cascades in the CeA. The crucial role of NDMA receptors is evidenced by the effect of NMDA-induced lesions in the PnC in studies that have completely eliminated startle (Davis, 2006).

Fear extinction is a result of repeatedly encountering a previously conditioned CS+ without the previously paired UCS. When presented with the CS+ alone, new learning occurs. Specifically, fear extinction involves the learning of a non-aversive representation of the CS+; acquisition, consolidation, and expression of this new non-aversive representation of the CS+ is thought to occur through a complex system involving the amygdala, prefrontal cortex,

hippocampus, and brain stem, among other areas (Davis, 2006). Mounting evidence suggests that extinction is mediated through inhibitory signals from cortical regions that act on the amygdala to inhibit expression of the previously acquired fear conditioning (Norrholm, 2015). In contrast to findings in studies of fear acquisition, rodent studies of NMDA receptor antagonism of the CeA and vmPFC as means of impairing fear extinction learning have found null results (Zimmerman & Maren, 2010; Lebron, Milad, & Quirk, 2004). Studies have found significant impairment when manipulating the following areas: basolateral amygdala (Zimmerman & Maren, 2010), nucleus accumbens (Holtzman-Assif, Laurent, & Westbrook, 2010), mediodorsal thalamus (Lee et al., 2012), and lateral amygdala (Kim et al., 2015). NMDA receptors in the amygdala seem to be crucial in extinction of conditioned fear as pharmacologically enhancing these receptors augments extinction while blocking them during or after extinction training blocks extinction, suggesting that NMDA receptors are important for the acquisition and consolidation of extinction (Davis, 2006).

Though there is overlap, neurobiological aspects of fear extinction recall differ from fear extinction learning. For example, the ventral medial prefrontal cortex (vmPFC), which is activated in both fear extinction learning and recall of fear extinction, may play a crucial role in extinction recall, but not extinction learning. Lesions to this region impair extinction recall in rats one day after conditioning, but do not impair extinction learning immediately after conditioning (Quirk, Russo, Barron, & Lebron, 2000; Lebron et al., 2004). Though the neural circuits involved in fear extinction recall are less thoroughly explicated than those involved in fear acquisition and fear extinction learning, existing evidence does show that fear extinction recall is a distinct aspect of fear learning.

Taken as a whole, these findings demonstrate that fear learning paradigms can index deficiencies in several separate neurobiological processes that are critical aspects of fear acquisition and extinction. In order to obtain a holistic picture of differences in extinction learning, researchers must consider processes of fear extinction recall. The paradigm used in this current study accounts for each of these aspects.

1.4 Fear Learning and RDoC

One aim of RDoC is to provide a framework within which researchers may formulate and evaluate intermediate phenotypes that may be useful in elucidating biological underpinnings of clinical problems. Intermediate phenotypes have been defined by three characteristics: (a) biological elaboration, (b) connection to clinical difficulties, (c) and modeling in animal and human studies (Briscione, Jovanovic, & Norrholm, 2014).

So far, studies examining the relationship between fear learning and psychopathology have employed central tendency statistics, which make group comparisons based on differences and/or change in arithmetic mean between groups and/or over time. Generally, studies of fear learning compare means between two groups (e.g., PTSD vs. no-PTSD) using ANOVA (e.g. Norrholm et al., 2015; Mineka & Oehlberg, 2008). Analyses using central tendency statistics do not account for the possibility that latent sub-populations of fear learning trajectories may exist within these groups (Galatzer-Levy et al., 2013b). Thus, they may not provide an accurate picture of the true relationship between fear learning and the clinical difficulties the researcher wishes to examine.

Recent studies have shown that Growth Mixture Modeling (GMM), which is a well validated method for detecting heterogeneity in longitudinal growth patterns of symptom change

over time, may be applicable to fear conditioning studies. In 2013(b), Galatzer-Levy and colleagues applied Latent Class Growth Analysis (LCGA), which is a simplified GMM that fixes within-class variability in growth parameters to zero (i.e. assumes a common intercept and slope for all members of each class), to archival data from a previous study investigating fear conditioning in rodents (n=58), in which fear was operationalized as freezing behavior during a one day fear learning paradigm; they discovered evidence of previously overlooked heterogeneity in fear extinction trajectories in this rodent population. Specifically, this analysis detected a "rapid extinction" group (57.3% of the sample) which extinguished fear quickly, a "slow extinction" group (32.3% of the sample) which extinguished fear after a delay, and a "failure to extinguish" group (10.3% of the sample) which did not extinguish fear.

Very recently, the findings from this original rodent model were translated to a human study. In 2017, Galatzer-Levy and colleagues applied Latent Growth Mixture Modeling (LGMM), which is a form of GMM that allows for within-group variations in growth parameters and estimates variance and covariance of slope and intercept for each class, to archival data pooled from several studies employing a common fear learning paradigm. This 2017 manuscript included two studies: study 1 examined LGMM classes in humans (n = 724) and operationalized fear responses using fear potentiated startle (FPS) during a one day fear conditioning paradigm; study 2 examined LCGA classes in a sample of mice (n = 127) which operationalized fear responses using freezing behavior during a two day fear learning paradigm (i.e. fear extinction recall was tested on day 2). Consistent with the 2013 study, results from the LCGA of rodent freezing behavior showed three latent groups based on differences during fear extinction, all three of which acquired the fear: "non-extinguishers" (30% of the sample), "slow extinguishers" (45% of the sample), and "rapid extinguishers" (25% of the sample). During fear extinction

recall (day 2 testing), the "non-extinguishers" exhibited the highest initial and final levels of freezing behavior. The human aspect of the study also exhibited three latent groups. However, unlike rodent models, the majority of humans did not acquire the fear. Among those who did, one group extinguished the acquired fear (i.e. evidenced a statistically significant negative slope during extinction trials) and one did not. Thus, in humans, LGMM trajectories were as follows: "Modal FPS Responders" (78.9% of the sample) who did not acquire the fear, "High FPS Extinguishers" (14.8% of the sample) who acquired the fear and extinguished it, and "High FPS Non-Extinguishers" who acquired the fear and did not extinguish it.

1.5 Fear Learning Methods

Though the aforementioned studies demonstrate that LGMM and LCGA are promising methodological approaches to elucidating fear learning phenotypes, they only represent initial steps toward validating this approach. Fear learning paradigms are very specific procedures; subtle differences may lead to dramatically different results. Prior studies utilizing more traditional central tendency-based analyses have shown that subtle differences in experimental procedures may lead to different statistical outcomes (Sehlmeyer, Schoning, Zwitserlood, Pfeiderer, Kircher, Arolt, & Konrad, 2009). The impact of methodological changes on expected statistical outcomes appears to stem from differences in the ways in which various experimental procedures may stimulate fear circuitry (e.g. Yuzhe, Nakae, Ishii, & Naoki, 2016; Chunm-hui, Berke, & Maren, 2010). This is important because these experimental differences may reflect naturalistic differences in the ways in which variations in lived experiences may stimulate fear circuitry. Thus, these experimental differences may have substantive implications that should be considered when attempting to translate basic science research into clinical applications. A major

impetus for the application of GMM methods (i.e. LGMM and LCGA) to fear learning lab studies is that these analyses may eventually translate into increased understanding of practical clinical problems. Since the experiences which lead to clinical difficulties can never be as uniform as a laboratory study, it is important to examine the degree to which these findings may vary or remain consistent across different methodologies. High consistency across experimental methodologies would suggest that these latent phenotypes are robust and may be highly applicable to a range of clinical presentations. However, if they vary considerably in response to subtle differences in methodology, it will be important to eventually isolate and replicate specific responses to specific changes in methodology as well as to examine potential naturalistic analogs to these methodological changes. Specific methodological differences in fear learning studies that may be of interest to trauma researchers include: differences in the specific unconditioned stimulus (UCS) used, differences in rates of UCS pairing consistency during conditioning, and differences in the amount of time that passes between fear acquisition and extinction.

Though use of a standardized UCS across different experiments can increase confidence in findings, cost and safety concerns may present a major barrier to utilizing some UCSs. The most commonly used stimulus in fear learning studies has historically been electric shock, which is expensive, potentially dangerous when utilized by inexperienced researchers, and may not be approved in the IRB process at many institutions (Biopac). Recently, air puff has gained popularity, but this method may be too expensive or technologically challenging to install for some labs (HarvardApparatus). The use of the combination of an aversive scream and a fearful face as an unconditioned stimulus is safe, free, and available to any lab. This unconditioned stimulus has been evidenced to lead to fear acquisition in undergraduate populations, though it may not be as potent of a stimulus as electric shock or airblast (Glenn, Lieberman, & Hajcak,

2012). This relatively mild stimulus may represent a useful tool for labs that lack the resources or expert personnel to acquire and safely utilize shock or airpuff. However, since it is a relatively mild stimulus, its effectiveness in capturing differences in fear learning in a GMM framework is untested. Thus, this study is the first to examine the use of an aversive scream combined with a fearful face as an unconditioned stimulus in the context of fear learning.

In lived experience and in the lab, the consistency of the pairing of a naturally aversive UCS with an initially neutral stimulus may vary; variations in consistency are shown to impact learning processes. Though many fear learning lab studies utilize 100% pairing consistency during acquisition, aversive stimuli in the real world are rarely paired 100% of the time (Foa, E.B., Zinbarg, R., & Rothbaum, B.O., 1992). Lab studies have demonstrated that varying degrees of pairing consistency during acquisition lead to various rates of extinction. Specifically, a study which compared three conditions (100% consistency, 75% consistency followed by 100% consistency, and 100% consistency followed by 75% consistency) found that 75% consistency followed by 100% consistency led to robust fear acquisition and significantly delayed extinction in humans relative to the other two conditions (Grady, Bowen, Hyde, Totschi, & Knight, 2016). A recent computational modeling study which examined activity of specific neurons in mice demonstrated that activity in neural units associated with extinction is diminished during early extinction phases when pairing consistency is partial, in comparison to when it is complete (Yuzhe, Nakae, Ishii, & Naoki, 2016). Thus, it is important to extend this methodology to examine various rates of pairing consistency. Since the proposed UCS is relatively mild and individuals who evidence difficulties during extinction are relatively rare, it makes sense to attempt to extend extinction learning. Thus, this study utilized 75% pairing consistency during the first two blocks of acquisition followed by 100% pairing consistency in the latter two.

Finally, the amount of time that passes in between fear acquisition and the opportunity for extinction learning varies in real world settings. This is shown to impact fear extinction recall in the lab, which is theorized to mirror real world fear extinction recall. Specifically, immediate extinction training is shown to correlate with decreased fear extinction recall in humans and animals (Merz, Hamacher-Dang, & Wolf, 2016). Thus, this study tested an immediate extinction paradigm in order to increase the likelihood of identifying difficulties in extinction learning recall.

An additional consideration in the process of testing GMM's usefulness as an RDoCian analytic technique is the degree to which this approach applies across diverse samples. A strength of the aforementioned 2017 study by Galatzer-Levy and colleagues is that it investigated a generally understudied population by analyzing a pooled sample that was predominantly African American (75.13%) [followed by Caucasian (16.6%), and other ethnicity (8.27%)]. The sample had relatively equal proportions by gender (47.18% male), a mean age of 38.79 years (SD 11.36), and a mixture of traumatized individuals with PTSD and/or Major Depressive Disorder (MDD), traumatized individuals who did not meet criteria for PTSD and/or MDD, and healthy non-traumatized healthy controls. Though clinical information was not reported in all studies from which these data were pooled, regarding clinical characteristics that were reported, the study's pooled sample characteristics were as follows: Traumatized healthy controls (n= 223); PTSD (n=181); MDD (n=39); PTSD and MDD comorbid (n=22); non-traumatized healthy controls (n=131) (Fani et al., 2012, 2015; Jovanovic et al., 2006, 2009, 2010a, 2010b, 2012; Norrholm et al., 2011; Norrholm et al., 2006; Sawamura et al., 2016). As detailed below in the sample demographics section of results, this study examines a sample with different characteristics across numerous aspects of sample composition.

2.0 Hypotheses

The objective of this study was to further explore profiles of fear acquisition and extinction and extend this analysis to investigate fear extinction recall in humans using GMM for the first time. Since this thesis was originally designed and proposed prior to the publication of the first study to apply GMM analyses to a human sample, original hypotheses and modified hypotheses are presented. At the time of the thesis proposal, prior to the publication of the Galatzer-Levy et al. (2017) study, the following hypotheses were offered at the original thesis proposal meeting:

1) First, it was hypothesized that profiles mirroring those found in rodents (rapid extinction, delayed extinction, no extinction) would be supported by LCGA in a human sample.
2) Second, it was hypothesized that profiles of fear extinction recall would be supported by LCGA in a human sample and would differ from those found in extinction learning. This hypothesis was preliminary and not required to pass the thesis, given the challenges associated with collecting sufficient data at two time points.

In light of recent findings by Galatzer-Levy and colleagues, the following updated hypotheses were offered prior to data analysis:

1) First, it was hypothesized that profiles mirroring those found in the 2017 study would be supported by LCGA results in the current sample: "Modal FPS Responders" who do not acquire the fear, "High FPS Extinguishers" who acquire the fear and quickly extinguish, and "High FPS Non-Extinguishers" who acquire the fear and do not extinguish it.

2) Second, it was hypothesized that LCGA results would support a model with multiple distinct classes of profiles of fear extinction recall in this sample, though these may differ from those found in extinction learning.

3.0 Method

3.1 Demographics and Self-Report Data

Participants were undergraduate students at Virginia Tech. They were recruited electronically on the university's online Sona system, through which students can earn class credit by completing the survey. Participants completed a number of surveys for this study as well as additional measures that were not of interest in the current study. More specifically, participants completed a number of items pertaining to demographic variables such as age and race. They also completed a number of psychometrically validated self-report surveys; psychometrically validated measures of interest to this study are explained below.

The Trauma History Screen (THS) is a brief, 14-item self-report measure that inquires about 13 specific traumatic events that one may experience as well as any "other" events that may be traumatic; the THS also indexes the number of times events occurred and which aspect of the event bothered the respondent the most. The THS is shown to have good reliability and validity and compares favorably to many longer measures of trauma exposure (Carlson et al., 2011). The PTSD Checklist for DSM-5 (PCL-5) is a 20 item self-report questionnaire designed to assess for Posttraumatic Stress Disorder (Weathers et al., 2013). The PCL-5 has been empirically validated in numerous samples, including non-clinical samples (Ashbaugh et al., 2016). The Obsessive Compulsive Inventory (OCI) is a 40 item self-report questionnaire designed to assess for Obsessive Compulsive Disorder (Foa et al., 1998). The OCI provides an

adequate index of OCD, including non-clinical samples (Simonds et al., 2000). The Beck Anxiety Inventory (BAI) is a 21 item self-report questionnaire designed to assess for Generalized Anxiety Disorder (Beck, 1993).The BAI has been empirically validated in numerous samples, including non-clinical samples (Creamer et al., 1995).

3.2 Lab Participants

Students who completed the self-report survey were eligible to sign up for the lab portion of the study through Sona for additional extra credit. Seventy-eight students participated in this study on day 1, 55 of which returned for day 2. In order to increase feasibility, participants were able to return for day 2 on any day they could, as long as it was more than 24 hours after day 1. Of the 55 who returned, the mean amount of days in between day 1 and day 2 was 5 (min = 1, max = 29, SD = 6.41). Of those participants, data on two were lost due to a computer crashing, an additional five were removed from analysis due to being non-responders (i.e. did not evidence reflexive blink responses to any startle presentations), and four were removed due to excess noise. Thus, 67 participants are included in analyses, 48 of which participated in both days of the study. Exclusion criteria included individuals with active psychosis, major heart conditions, an inability to detect tones at 30 dB(A)SPL at frequencies ranging from 250-4000Hz. Additionally, research assistants conducted a brief in-person self-report survey in the lab in order to measure substance abuse on the day of the study. Specifically, participants were asked whether or not they had used: caffeine within the past 12 hours, alcohol within the past 24 hours, marijuana within the past 24 hours, or any other drugs within the past 24 hours. Since no recruited participants met any of these criteria, no participants were excluded due to these reasons. For a participant characteristics, see figure 1.

3.3 Apparatus

Lab testing was conducted in the Virginia Tech Mind-Body lab. In order to index startle, eye-muscle movement was assessed with EMG signals captured by two 5-mm Ag/AgCl electrodes, which are filled with electrolyte gel. Though they were not of interest to the present study, Heart Rate Variability (HRV) and respiration signals were collected for other analyses, which are outside of the scope of the present study. HRV data were quantified using Electrocardiogram (ECG) signals from three ECG electrodes; one ECG electrode was placed beneath each clavicle and one was placed on the participant's left rib. Respiration was quantified using a chest-respiration belt, which was fastened using Velcro and was adjusted to be "snug but comfortable." The EMG and ECG signals were amplified and then digitally sampled by the Biopac MP150 data acquisition system. All physiological data were recorded and cleaned with Biopac's AcqKnowledge software on a PC. Stimulus presentation was coded and presented using DMDX software, which is a Windows-based program designed for experimental stimulus presentation; DMDX uses the features of Pentium class computers to provide accurate timing and synchronization of visual and audio output (Forster & Forster, 2003).

3.4 Stimuli

The startle probe was a 105dB, 50-ms burst of broadband noise. Conditioned stimuli were two different female faces drawn from the NimStim, which is research-based collection of standardized facial expressions being displayed by various actors which are shown to reliably depict specific emotions that are easily recognizable for study participants (Tottenham et al., 2008). The unconditioned stimulus was a loud shrill sound of a girl screaming at 98 dB for 2.4 seconds paired with one of two fearful female faces chosen from NimStim. The specific faces used were NimStim F01 and NimStim F03 (Macbrain). The assignment of F01 or F03 to act as

the CS+ or CS- was counterbalanced across study participants. Both CS+ and CS- employed a neutral, closed mouth expression (NC); only the CS+ employed an open mouth fearful expression (FO), which was presented concurrently with the loud shrill sound of a girl screaming immediately after the presentation of the neutral closed mouth expression. Thus, the CS+ consisted of the presentation of a picture of a specific female displaying a closed mouth neutral facial expression which predicted the concurrent presentation of a loud shrill sound of a female screaming displayed concurrently with a picture of that same female displaying an open mouth fearful facial expression. Noises were played through noise-cancelling Bose headphones. The CS- consisted of a presentation of a different female displaying a neutral closed mouth expression which never predicted the presentation of an open mouth fearful face or a loud shrill sound of a female screaming. Stimulus presentation of CS+ and CS- is depicted in Figure 3.

The use of the above-mentioned combination of a girl screaming and a fearful female face has been tested and shown to condition fear in healthy college samples, though overall responding was less than electric shock and it was not as effective in discriminating individuals with symptoms of anxiety (Glenn, Lieberman, and Hajcak, 2015). Currently, the degree to which this CS+ is useful in explicating patterns of fear conditioning is unclear. Thus, the current study adds additional evidence from which labs which may inform future studies which may consider employing this CS+.

3.5 Conditioning Paradigm

This study employed a two day conditional discrimination paradigm (see figure 2). Day 1 included a brief acclimation period, followed by a fear acquisition phase, followed by a fear extinction phase. There were no breaks between phases and the entire session was an average of

45 minutes in length. After the session, subjects were asked to rate the averseness of the scream and the startle probe on a scale from zero to nine.

The conditioning paradigm began with an acclimation period consisting of four startle probes presented along with four female faces from NimStim, each of which depicted a closed mouth neutral expression and none of which were used as CS+ or CS-. Specifically, NimStim 05F, 06F, 07F, and 10F were displayed sequentially. Then, the fear acquisition phase began immediately after the acclimation period. Fear acquisition consisted of four blocks of 4 presentations of CS+ and CS- (startle probes occurred on 75% of trials) as well as 3 inter-trial interval (ITI) startles for a total of 16 trials, 32 ITIs, and 27 startle probes; CS-UCS pairing consistency was for the first 2 blocks and then 100% for the second 2 blocks. The fear extinction phase immediately followed the fear acquisition phase and consisted of eight blocks of 4 presentations of CS+ and CS- (startle probes occurred on 75% of trials) as well as 3 ITI startles for a total of 32 trials, 64 ITIs, and 54 startle probes; all blocks of fear extinction consisted of a 0% pairing consistency (i.e. fear extinction was conditioned via removal of the UCS). Fear extinction recall was tested on a separate day at least 24 hours after fear extinction. Due to the logistical difficulties associated with recruiting participants at two timepoints, participants were allowed to come in at any time beyond 24 hours. The specific amount of time between day 1 and day 2 was recorded. The day 2 trials and consisted of eight blocks of 4 presentations of CS+ and CS- (startle probes occurred on 75% of trials) as well as 3 inter-trial interval (ITI) startles for a total of 32 trials, 64 ITIs, and 54 startle probes; all blocks of fear extinction consisted of a 0% pairing consistency (i.e. fear extinction was conditioned via removal of the UCS).

All blocks included: 4 CS+, 4 CS-, and an inter-trial interval (ITI) in between each stimulus presentation (see figure 3 for trial structure). ITIs lasted for a duration of 10-12 seconds with a startle a either occurring (3 times per block) or not occurring (5 times per block) anywhere from 4.5 to 6.5 seconds into the ITI. ITI length and timing of startle probe were both randomized. Conditioned stimulus presentation consisted of the presentation of either one or two pictures of the same face from NimStim. The neutral closed mouth facial expression was displayed for 7.5 – 8 seconds with startles occurring (75% of time) at 3.5-4.5 seconds, while the fearful open mouth expression was displayed immediately after the neutral closed mouth facial expression (during 75% of CS+ trials in early acquisition and 100% of CS+ trials in late acquisition) and was displayed for 2.4 seconds. The details of stimulus presentation order and trial and ITI structure are depicted in Figure 2 and Figure 3, respectively. Consistent with recommendations from fear learning methodological literature, the order of trial presentation and the specific length of trial presentation and ITI within aforementioned time intervals was pseudo-randomized (i.e. randomized but with the following rules: no more than 2 same CSs in a row; standardized blocks of 8 presentations; 4 each of each CS with 3 of the 4 of each being a startle; 3 ITI startles and 5 without startle) (e.g. Glenn et al., 2012). Pseudorandomization consisted of 2 pseudorandomized orders of CS+/CS- presentation and trial timing, each of which included two versions (one with NimStim F01 as the CS+ and NimStim F03 as the CS- and one with F03 as the CS+ and F01 as the CS-). Thus, there were four total trial presentations, all of which followed the same aforementioned structure in terms of blocks and trial design.

3.6 Procedure

Participants were recruited electronically on the university's online Sona system, through which students can earn class credit by completing the survey. The portal and web-link in Sona brought participants to an online survey. Participants were then presented with confidentiality information and asked whether they understand their rights as a participant. After confirming that they comprehended these rights, they were able to complete the survey, which included a variety of mental health measures which are part of a larger study. Upon completing the survey, participants were invited to sign up for timeslots for this study, which took place in Williams Hall.

After being recruitment and upon arrival to the laboratory, Room 251 in Williams Hall, a member of the research team greeted participants and sat them in front of a computer. The researcher then orally explained the procedures, purpose, benefits, and risks associated with the study. Then, the experimenter asked participants to summarize the study procedure, after which participants were asked to read and sign an informed consent sheet.

Following written and oral consent, three disposable adhesive electrodes were placed on the participant's face (two electrodes were below the right eye and one was be on the forehead) for facial electromyography (EMG); and another two electrodes were placed on the participant's right and left clavicles and one in the abdominal region for electrocardiography (ECG). Also, a band containing a respiratory transducer was attached around the chest. The electrodes and the band were attached on the participant by a gender-matched research team member. From this point, facial EMG, ECG, and respiratory signals were recorded from the individual throughout the study session.

Hearing impairments were examined by requiring participants to detect tones at 30 dB[A]SPL at frequencies ranging from 250 to 4000 Hz. These assessments were used to ensure eligibility for the study. After these assessments were completed, a member of the research team entered the lab and explained the remainder of study instructions. Participants were first asked to sit quietly for 5 minutes while baseline cardiovascular and respiration data were collected, which are to be used in separate analyses which are beyond the scope of this thesis. Then, the researcher loaded the DMDX program, which presented study stimuli and instructed participants to begin the program once prepared, and then to exit the room. All further experimental instructions and stimuli were presented via DMDX software on the lab PC. Researchers then observed from a nearby room and monitored the amount of distress the participant was in as well as monitored to make sure that the software program was running correctly.

Participants were invited to return for a second lab day to test fear extinction recall, which followed the same protocol as mentioned above, beginning with paragraph three of this subsection of this document (subsection 3.5 procedures).

Before the study began, all procedures were approved by the Virginia Tech Institutional Review Board in order to assure participant safety and proper adherence to ethical guidelines.

3.7 Data Reduction

Startle responses were quantified using EMG which were acquired with BIOPAC MP 150 (Biopac Systems, Inc., Aero Camino, CA). Once acquired, data were filtered, rectified, and smoothed based on all procedural recommendations from Blumenthal et al., 2005 using Acqknowledge software (Biopac Systems, Inc., Aero Camino, CA). Once filtered, rectified, and smoothed, startle responses were linked with startle probes. Then, each individual startle probe

for each participant (n=11,027) was examined for validity; invalid blinks (i.e. blinks in which there was excess noise, blinks which began before the startle probe, or trials in which a spontaneous blink occurred immediately before the startle probe) were deleted based on recommendations from Blumental et al. (2005). Specifically, 5.79% of individual startles were deleted (n=638). Additionally, participants who did not respond to any startle probes (n=5) were eliminated from analysis (Blumenthal et al., 2005). Then, data were exported and cleaned using SPSS (IBM Corp). In SPSS, a Mean Score was first created for each CS type as well as for ITI within the block (mean score = average magnitude of startle reaction to all three startle probes for each CS and ITI). Next, a Difference Score for each CS type was created by subtracting the mean score startle magnitude to the ITI noise probe from startle magnitude in the presence of a CS in each stimulus presentation. Finally, data were once again exported and analyzed using Mplus 6.11 and Mplus 8.0 (Muthén & Muthén, 1998-2012). Difference scores were used for all analyses.

3.8 Analysis: Demographics and ANOVA

First, demographic and self-report statistics were analyzed in SPSS (see figure 1). Self-report psychological symptoms were indexed via psychometrically validated measures, which are specified in the results section. Caffeine, drug, and alcohol use were also collected via verbal report from each participant before completing the experiment. Potential impacts of substance use were explored using ANOVA with the within subjects factors of CS (2 levels) and Fear Learning Phase (3 levels day 1; 2 levels day 2) and between subjects factor of caffeine use.

Next, a manipulation check was performed using repeated measures ANOVA in SPSS. In order to account for the loss of subjects in day 2 of the experiment, an ANOVA model was run

for day 1, a separate model was run for day 2, and a final model was run for both days. Day 1 and day 2 models included the within-subjects factors of block (12 levels Day 1; 8 levels Day 2) and CS (2 levels) as well as the between-subjects factor of Trial Order (4 levels). The ANOVA for both days substituted block with phase (5 levels). Based on recommendations from Van Ginkel & Kroonenber (2014), multiple imputation was performed to deal with all randomly missing data (n = 638 randomly missing startle trials). Prior to conducting ANOVA, averages were computed for CS+ and CS- trials for each block, which are presented in Figure 5.

3.9 Analysis: LCGA

Latent groups of fear acquisition and fear extinction were tested using LCGA in Mplus (Muthén & Muthén, 1998-2012). For the current study, estimates were obtained using maximum likelihood (ML) via expectation-maximization (EM) (Ram & Grimm, 2013). Since Mplus uses full information maximum likelihood estimation, missing data were not imputed for LCGA. Thus, missing startle trials were not included in LCGA analysis (Muthén & Muthén, 1998-2012).

Prior to comparing competing models, models were examined for potential estimation problems that could lead to biased or untrustworthy results. Specifically, tests were run to check for issues pertaining to non-normal probability distribution, local maxima, and model non-identification due to inappropriate data. Univariate skewness and kurtosis was examined using SPSS descriptive statistics while multivariate skewness and kurtosis comparisons were derived from the SK test in Mplus (IBM; Muthen & Muthen, 1998-2011). Based on recommendations from Muthen and Asparouhov (2015), models using the skewed t-distribution were tested and compared to normal GMM models based on both statistical fit and theory. Global solutions were assured by examining and replicating the maximum likelihood estimation (MLE). Non-

converging models were first re-run using an increased number of starting values and MLE iterations. Consistent with recommendations from Wickrama and colleagues (2016), models that did not converge after increasing starts to 1000 and iterations to 50 were removed from consideration.

Competing models were compared based on a holistic evaluation. First, models were examined based on information criteria [Bayesian (BIC), sample-size adjusted Bayesian (SSBIC), Aikaike Information Criterion (AIC)], likelihood ratio tests [Lo-Mendel Rubin Likelihood Ratio Test (LMRLRT) and the Bootstrap Likelihood Ratio Test (BLRT)], and model fit (entropy). GMM models for which the smallest class has an n of less than 25 or which constitute less than 5% of the total sample were interpreted with extra caution. Additionally, parsimony and theory were weighed in all model decisions (Wickrama, 2016).

4.0 Results

4.1 Demographics, Self-Report, and Substance Use

Sample demographics and self-report scores for the participants in this study are listed in Figure 1. The study sample, as shown in Figure 1, evidences a gender distribution that is somewhat different than Virginia Tech's overall gender distribution (57.5% male) and is comparable to Virginia Tech's racial diversity, with the most notable difference being a lower Asian sample population than is found in the school on average (9.3%) (college factual). While the FSS indicates a high average number of Criterion A traumas reported among those reporting their number of traumatic experiences, the number is comparable to trauma exposure levels found in the general population (Kilpatrick et al., 2013). The average scores for self-report indices of fear and anxiety related clinical disorders and symptoms were all well below the

clinical cutoffs for those measures (Carlson et al., 2011; Weathers et al., 2013; Ashbaugh et al., 2016; Foa et al., 1998; Beck, 1993).

Self-report substance scores on the day of the lab study are also listed in Figure 1. Due to logistical considerations and the preliminary nature of this study, participants who used caffeine or other substances were included in the study. Since caffeine was the only substance used by a substantial proportion of the sample and caffeine use has been shown to influence acoustic startle response magnitude (though no extant studies specifically examine FPS), the potential effects of caffeine use in the study were examined (Benke et al., 2015). In order to examine the likely impact of caffeine use in this sample, ANOVAs were run for day 1 and day 2. Caffeine use had no significant effect anywhere in the day 1 model. Thus, the impact of caffeine use at day 1 is not a major concern. For day 2, caffeine use had a significant main effect on EMG startle scores ($p=.035$; $F=4.77$). Thus, it should be noted that caffeine may have impacted results during extinction, though no interactive effects were found with caffeine on day 2.

4.2 ANOVA for Manipulation Check

As hypothesized, the day 1 ANOVA results showed main effects of block on EMG difference score magnitude $F(11,53)=3.88, p<.001$ and CS $F(1,63)=46.885, p<.001$, which were due to higher overall responding to CS+ trials than to CS- as well as an overall decreases in mean score during extinction. Additionally, block*CS displayed a significant interaction effect $F(11,53)=2.10, p=.036$, which was indicative of higher overall discrimination between CS+ and CS- during acquisition. This indicated that some degree of discriminant learning seemed to occur and strengthen over trials. However, an overall decreasing trend for CS+ difference score magnitude began during acquisition and may have indicated habituation in the overall sample.

Trial Order did not display significant main effects or interactions, indicating that pseudo randomization was successful and that trial order as well as which specific face was used as a CS+ did not impact results. For day 1 ANOVA results, see Figure 4.

As hypothesized, the day 2 ANOVA showed a significant main effect of CS type on EMG difference score magnitude $F(1, 44)=9.68$, $p=.003$. Surprisingly, this was due to a higher mean score of the CS- in comparison to the CS+. This ANOVA also showed two significant interaction effects: CS * Trial Order $F(3,44)=7.74$, $p<.001$, as well as Block*CS* Trial Order $F(21,110)=2.07$, $p=.008$. The CS * Trial Order interaction was driven by CS+ scores being higher for Trial Order 1 only. The Block*CS* Trial Order interaction seemed to derive from occasional higher magnitudes for the CS- relative the CS+ which only occurred in specific Trial Orders during specific blocks. Specifically, the CS+ magnitude was higher than the CS- magnitude in the following: Block 1 (Trial Order 1), Block 2 (Trial Order 1, 2, 3), Block 3 (none), Block 4 (Trial Order 1), Block 5 (Trial Order 1 and 4), Block 6 (Trial Order 1 and 2), Block 7 (Trial Order 1), Block 8 (Trial Order 4).

Day 2 manipulation check results should be interpreted with caution. In repeated measures ANOVAs, Type I error rates are often inflated and can be particularly problematic in models using lower sample sizes (Park, Cho, & Chang-Seok, 2009). Thus, when a finding does not align with theory, findings from other studies, or common sense, it is possible that Type I error is present. The finding that the overall responses to CS- were higher during Day 2 is counterintuitive and no extant studies report enhanced CS- responding during fear extinction recall, while many do report enhanced CS+ responding in comparison to CS- (e.g. Straus et al., 2017;). Thus, due to a lack of theoretical support for this finding as well as the abundance of

studies which find the opposite, this finding should be interpreted cautiously as it may be a result of Type I error. The finding of several interaction effects for trial order is also best explained as likely Type I error. Though trial order repeatedly emerged with interaction effects, there was no consistent pattern in these interactions. Furthermore, trial order did not differ in any systematic way and participants were randomly assigned to groups. Overall, since the day 2 ANOVA results did not align with theory or common sense, were conducted in a smaller sample than those in the day 1 study, caffeine was identified as a potentially influential variable at day 2 in a separate analysis, and the days in between lab visits for day 2 data collection varied considerably, the true reason for these findings is unclear. Overall, these findings indicate that, unlike day 1 data, day 2 data may not have derived from a successful manipulation and should be interpreted with caution. For day 2 ANOVA results, see Figure 5.

As hypothesized, the ANOVA for the entire experiment showed a significant main effect of phase on EMG difference score magnitude $F(4,41)=11.24, p<.001$ and CS $F(1,44)=5.05, p=.03$. As hypothesized, main effects for CS were driven by increased responding to the CS+ and main effects for phase were driven by decreasing overall responding throughout day 1 (highest during acquisition, lowest during late extinction) but increasing overall responding from late extinction to early recall as well as from early recall to late recall. It also showed interaction effects for Phase* Trial Order $F(12,108)=2.28, p=.01$, and Phase*CS $F(4,41)=4.83, p=.003$. Interaction effects for phase*CS were driven primarily by a rapid decrease in CS+ responding during late extinction. Interaction effects for phase* Trial Order were driven by a sharp increase in late recall responding by Trial Order group 3 in comparison to other Trial Order groups. Overall, there results align with findings from day 1 and day 2 ANOVAs. For ANOVA results from models employing data from both days, see Figure 6.

4.3 LCGA - Addressing Estimation Problems

In these data, univariate tests of skewness and kurtosis indicated non-normality in a number of variables. Specifically, positive skewness was found in a total of five blocks, negative skewness was found in one block, and positive kurtosis was found in a majority of blocks. Additionally, SK test results showed that multivariate skewness and kurtosis for extracted classes differed from the sample in every normal GMM model run. Thus, Skewt distribution models were run and compared with normal GMM models based on model fit and theoretical interpretability.

Local maxima were generally not an issue in these analyses. Thus, solutions that were offered for models that converged successfully were global solutions. The only model which may have returned a local maxima was the three class model for the two day piecewise model, which was further disqualified from consideration based on non-convergence and low sample size in the smallest class. Thus, local maxima did not influence overall interpretation of findings and are not a major concern in this study.

All day 1 models successfully converged. Thus, model non-identification and inappropriate data were not a major concern for acquisition and extinction aspects of this study. However, piecewise models attempting to model separate slopes for the entirety of all three experimental phases did not converge when a quadratic parameter was included, even after reducing fear recall blocks to the minimum needed to model an additional slope. Skewt models testing more than two classes did not converge, with the lone exception of a three class model for acquisition. Additionally, linear plus quadratic Skewt models did not converge and Skewt models for acquisition, extinction, and recall did not converge. Thus, these models were not interpretable.

4.4 LCGA - Model Fits and Parameters

First, model fit was compared for a one through three class linear change model which included all four blocks of acquisition (see Figure 7 for comparative fit indices). Information criteria decreased and entropy increased with each successive model. However, the two and three class models provided comparable values for BIC and entropy. Furthermore, the Lo-Mendell Rubin Likelihood Ratio Test (LMR-LRT) for the two-class model trended toward significance for the two class solution (p=.058) but not for the three-class model (p=.27). While the Bootstrap Liklihood Ratio Test (BLRT) was significant for each, the three-class model's BLRT likelihood value did not replicate, indicating that the BLRT may not be trustworthy for that model.

Next, class composition information revealed that the three class model's smallest class comprised just 3.1% of the sample (n=2). Thus, the three class solution was rejected and the two class solution was selected (see figure 8 for parameter estimates and Figure 10 for a graph). This solution included an acquiring class (n=7; 12.5% of sample) with a positive slope (slope = 0.177; p=.002) and a habituating class (n=60; 87.5%) with a negative slope (slope=-0.061; p=.004). Models were run with the intercept set at block 1 and block 4 in order to provide a sense of overall increase; all other estimates were the same for both models. The acquiring class evidenced a significant intercept at block 1 of acquisition (intercept = 0.282, p<.001) and at block 4 (intercept = 0.814, p<.001). The habituating group evidenced a significant intercept at block 1 of acquisition (intercept = 0.242, p<.001) but not at block 4 of acquisition (intercept = 0.060, p=.092).

Model fits were also compared for a one through four class linear plus quadratic model for acquisition. The Bayesian Information Criteria (BIC) was lowest for the two class solution. Entropy was lowest for the three class solution, though the two class solution was comparable.

For all solutions, the LMR-LRTs were well above the p=.05 cutoff for significance and the BLRT failed to replicate. Slopes were not significant for any solution. Overall, this suggests a lack of reliable fit to the data. The two class solution fit 12.4% of the sample (n=7) into a group that evidenced an increase in FPS magnitude and 87.6% of the sample (n=60) into a group that evidenced a decrease in FPS magnitude. However, neither the slope nor quadratic function was statistically significant in either group. Overall, the two class linear model showed better fit statistics, is more parsimonious, and is a better fit with theory. Thus, the linear model for acquisition compared favorably to the linear plus quadratic model.

Models were examined for fear extinction recall. Since these models ignored all day 1 data and set the first block of recall as the intercept, they were interpreted with caution. For both linear only and linear plus quadratic models, two class solutions included just one participant in the smaller class; no further models were explored.

Piecewise models were explored for acquisition and extinction (see figure 7). Phases were connected by a single intercept. Though the experiment included 8 blocks of extinction trials, models including all trials evidenced local solutions and were thus uninterpretable. Thus, final models were run through the first four blocks of extinction only. In order to provide additional information regarding change in CS+ responding, analyses were run with the intercept placed at block 1 as well as at block 4 of acquisition; intercept placement did not impact estimation of other parameters. The three class linear solution was eliminated due to a smallest class with n=2 (3.2% of sample). Information criteria indicated improved model fit from a one to a two class model and entropy indicated high likelihood of correct placement into classes.

However, the LMR-LRT was not significant (p=.142) and the BLRT did not converge, indicating an unreliable BLRT.

The two class model was selected (see figure 8 for parameter estimates and figure 11 for a graph). Consistent with the linear model in acquisition only, this model indicated a small acquiring class (16.9% of the sample; n = 11) with a positive intercept when set at the first block of acquisition (intercept = 0.322; p = .022), a positive slope during acquisition (slope 1 = 0.12; p=.006), and a positive intercept when set at the fourth block of acquisition (intercept = 0.682; p<.001) as well as a larger habituating group (80.9% of the sample; n=56) with a positive intercept when set at the first block of acquisition (intercept = 0.233; p<.001), a negative slope during acquisition (slope 1 = -0.066; p=.001), and an intercept that was not significantly greater than zero when set at the fourth block of acquisition (I=.035; p=.063). This model indicated that the acquiring group successfully extinguished fear (slope 2 = -0.064; p=.004) and the habituating group evidenced no significant change in fearful responding during extinction (S2=0.00; p=.971).

Linear plus quadratic piecewise models were also tested for acquisition and extinction. As was the case in the linear plus quadratic models for acquisition only, these models evidenced two groups with overall trends that mirrored those found in the linear only models. Graphs and mean scores suggested a minority subgroup which had an increase in FPS during acquisition followed by a decrease in FPS during extinction and a majority group which had a decrease in FPS during acquisition and extinction. However, the growth parameters were not statistically significant and fit statistics were slightly worse than those in the linear only model. Thus, the more parsimonious and meaningful linear model was selected.

Models were tested which included acquisition, extinction, and extinction recall, beginning with linear only models. The linear only three class model was eliminated due to a small third class (6% of the sample; n=3). The two class model evidenced a similar overall pattern to the piecewise model from day 1, with the addition parallel flat lines during recall. Due to study dropout, day 2 models included a smaller sample size (n= 48). Thus, results should be interpreted carefully and this part of the analysis should be viewed as preliminary.

Due to significant SK tests, Skewt models were compared to normal models. All Skewt models produced at least one estimate of infinity for Skewness, Degrees of Freedom, or both. Thus, these models were interpreted with caution as the Skewt distributional assumption may not be appropriate for these data and the initial observed non-normality may have been a result of different classes, which may be best captured by normal models. Skewt models testing for acquisition converged for a one, two, and three class linear model. The Skewt model for acquisition demonstrated improving fit indices for the two class model, but results did not align with theory. More specifically, this model indicated a majority class (n=57; 85%) with a non-significant downward slope (slope = -.028; p=.098) and a significant negative intercept (I= -.01; p=.013) and a smaller subpopulation (n=10; 15%) with a non-significant downward slope (slope = -.001; p=.872) and a significant positive intercept (I=.462; p<.001). Similarly, piecewise models for acquisition and extinction converged for a two class model and showed improved fit indices compared to the one class solution, but results did not align with theory. More specifically, this model indicated a majority class (n=59; 87%) with a non-significant downward slope during acquisition (slope = -.031; p=.085), a significant negative intercept at the end of acquisition (intercept = -.03; p=.008), and a nonsignificant upward slope during extinction (slope 2 =.005; .331) as well as a smaller subpopulation (n=8; 13%) with a significant downward slope

during acquisition (slope = -.008; p=.845), a significant intercept at the end of acquisition (intercept =.462; p<.001), and a nonsignificant positive slope during extinction (.03; p=.06).

Though the over extraction of spurious classes is of concern in non-normal data such as these, simulation studies which suggest examining the skewt distribution for non-normal data also note that normal distributions have several noteworthy strengths in comparison to skewt models. Importantly, it is noted that normal mixtures may be more suitable in studies with sample sizes less than 100 and that normal mixtures should be given especially careful consideration when theory suggests that non-normality in the overall data may be caused by meaningful latent classes or that meaningful but small latent classes may exist (Muthen & Asparouhov, 2015). Evidence from the only known previous GMM study in humans suggests that a small minority are particularly susceptible to conditioned fear and a previous GMM rodent study suggests that the presence of heterogeneity in natural fear responses is the cause of the non-normality that is commonly seen in fear learning data (Galatzer-Levy et al., 2017; Galatzer-Levy et al., 2013 b.). Furthermore, simulation studies have shown that when normal GMM is incorrectly applied to non-normal data, the Bayesian Information Criteria (BIC) may decrease even when classes do not align with best practices regarding interpretation of GMM (e.g. when a class includes less than 1% of the population and does not align with theory). The mathematical explanation for this phenomenon is that the GMM assumption of within-class normality may be met by using additional classes to fit the distribution (e.g. a class may be formed simply to match a long tail in a skewed distribution) (Muthen & Asparouhov, 2015). However, in these data, the BIC for the piecewise linear normal LCGA increased when comparing a two class to a three class solution while the linear normal acquisition model only decreased minimally for the three class solution. Though the decreasing BIC in the two class model does not guarantee that the

normal LCGA fits these data, this increases confidence in these findings and, in conjunction with the small sample size and the existence of a small but meaningful latent class during acquisition, supports the theory that a normal LCGA may be appropriate for these data. For those reasons as well as the aforementioned superior theoretical explanatory value of the normal models, the best fitting normal models were interpreted and the skewt models were rejected.

Finally, models of FPS to the CS- were also explored in order to provide a more comprehensive picture. Models for CS- clearly indicated a one class solution as all two class solutions utilized a very small second class, including: the linear model of acquisition (n=1), the piecewise linear model of acquisition and extinction (n=2), the linear plus quadratic models for acquisition (n=1), and the linear plus quadratic model for acquisition and extinction (n=1).

4.5 Mean-Centered Analyses

One class base models of LCGA provide information about the group as a whole. For the model during acquisition, results indicate that the group average evidenced a significant positive intercept during block 1 (intercept = 0.255; $p<.001$), a non-significant slope during acquisition (slope = -.032; $p = .137$), and a significant positive intercept at block 4 (intercept = .159; $p<.001$). For the piecewise model, results indicate that the group average evidenced a significant positive intercept at block 1 (intercept = 0.224; $p<.001$), an insignificant negative slope during acquisition (slope = -.037; $p = .079$), a significant positive intercept at block 4 (intercept = 0.113; $p < .001$), and a non-significant negative slope during extinction (slope = -.011; $p=.073$). Thus, the sample as a whole indicated no significant change.

A post hoc one-way ANOVA with a within subjects factor of block examining the change in CS+ responding during acquisition trials found a significant effect

F(3,64)=3.34,p=.025. More specifically, this test indicated a decrease in overall CS+ responding with pairwise comparisons showing just one statistically significant change, which was a decrease in mean score of -.147 from block 2 to block 3 (p=.024). The difference in mean from the first to the final block of acquisition was -.055 (p=1.0). These results would suggest that the sample as a whole habituated to the CS+ and habituation primarily occurred between block 2 and block 3.

Comparisons of mean changes between blocks 1 and 4 of acquisition in the ANOVA (.055) with the estimated slope found in the habituating class from the LCGA model for acquisition (slope =-.061) and change in mean between blocks 1 and 4 of acquisition (change in intercept = .182) reveal that a majority of this population may habituate to a degree that is estimated to be 3.3 times as great as the overall group mean while a small but significant subpopulation may evidence fear acquisition rather than habituation.

5.0 Discussion

5.1 Summary and Synthesis of Results

As they pertain to hypotheses, the main finding of this study disconfirm hypothesis 1. Results in this investigation indicated a two class solution with a minority group that acquired the fear and then extinguished and a majority group that habituated. Thus, the models indicated one less class than hypothesized and did not support distinct phenotypes of fear extinction. Hypothesis 2, which was exploratory, was not supported and day 2 results were not interpreted as meaningful due to an inadequate sample size for proper analyses. More important than the confirmation or disconfirmation of hypotheses are the substantive implications of these findings, which are discussed below.

Though hypothesis 1 was rejected, the main theoretical finding of this study is the identification of two distinct phenotypes of acute threat ("fear") acquisition in a convenience sample of undergraduates with low overall psychopathology. More specifically, this study identified one small latent group of less than 20% of the sample which demonstrates the acquisition of fear as well as a larger group comprising of greater than 80% of the sample which habituates to the aversive stimulus. This finding is consistent with a growing trend in the literature which indicates heterogeneity in response to stress. A number of animal studies have noted that heterogeneity in rodent learning models may suggest variability in conditioned threat response (McEwen et al., 2012). Recently, GMM approaches have been shown to capture meaningful heterogeneity in both humans and rodents, though the degree to which these findings may translate across methodologies and populations is unclear (Galatzer-Levy et al., 2013 b.; Galatzer-Levy et al., 2017). Consistent with the only known human study applying GMM in a fear learning paradigm, the present investigation suggests the existence of a distinct phenotype of high susceptibility to conditioning fear. Consistent with an RDoCian approach, this phenotype was identified using a bottom-up empirically driven approach rather than a top-down examination of a-priori categories.

A second noteworthy finding of this study was that, contrary to hypotheses, just two phenotypes were identified in this sample: a highly condition-able phenotype and a habituating phenotype. In contrast to a previous study by Galatzer-Levy and colleagues (2017), analyses did not reveal distinct fear extinction phenotypes. The current study unveiled one group which acquired the fear before extinguishing and a larger group which habituated rather than acquiring the fear. The design of this study differed from the other known study of GMM in a human fear learning study with regard to a range of potentially influential factors including: the

unconditioned stimulus used, the rates of CS-UCS pairing during conditioning, and the amount of time which passed between fear acquisition and extinction. Though it is impossible to rule these possibilities out, it should be noted that the differences in pairing consistency were designed to slow the rate of extinction in order to increase variability during that phase; this design choice is backed by evidence (Grady, Bowen, Hyde, Totschi, & Knight, 2016). Additionally, the immediacy of extinction trials is not generally linked with increased ability to extinguish fear, but is rather linked with deficits in extinction retention (Merz, Hamacher-Dang, & Wolf, 2016). While it is possible that the relatively mild UCS led to aversive associations that were more easily extinguished, it seems more likely that this result was due the use of a smaller and healthier sample. Furthermore, this study employed LCGA rather than LGMM due to the smaller sample size. Additionally, model fit indices suggested that linear models fit these data best. Conversely, linear plus quadratic models emerged as models of best fit in the study by Galatzer-Levy and colleagues (2017). Due to the smaller sample size and simplified models used, it is very possible that the current study was underpowered to detect a third latent phenotype. Since the study by Galatzer-Levy and colleagues (2017) found that just over 5% of their sample failed to extinguish fear, it seems unlikely that this study would have detected such a small class if it were present in the sample. Additionally, it is possible that the lack of a non-extinguishing group was a result of a healthier sample, a milder UCS, or a lack of time for consolidation of threat acquisition learning.

Analyses of CS- trials clearly indicate that a one class model best fit the data. Importantly, CS- trial data were highly non-normal. If the phenotypes identified for CS+ were a result of non-normality, one would expect CS- analyses to also identify spurious classes. Thus, the observed heterogeneity in CS+ trials is likely due to substantive differences in fear learning

rather than an extraction of spurious classes due to non-normality. Analyses of models run with a skewt distribution produced skewness parameters of infinity, suggesting that the distribution may not fit the data, and extracted theoretically meaningless classes. In addition to increasing confidence in results, this also suggests that other studies which seek to explore non-normal data using a GMM should explore normal GMM models alongside non-normal models.

Several other findings emerged which pertained to issues of study design. Though LCGA analyses of fear extinction recall did not provide theoretically useful information, the failure to confirm the second hypothesis does demonstrate the importance of having a large sample size for analyses of more than two phases of learning. Results from a manipulation check ANOVA indicated that the CS+ was effective overall and LCGA analyses showed that a subgroup of participants acquired fear to the stimulus. This demonstrates that a scream presented concurrently with a fearful female face may be a useful unconditioned stimulus for fear learning studies in relatively healthy college populations. A primary goal of this study was to investigate a statistical method that may provide information that examination of group averages would miss. Examination of manipulation check, post-hoc ANOVA results, and estimates from one class LCGA base models demonstrate the contrast between analyses of group averages and techniques such as GMM which explore heterogeneity. The stark contrast highlights the need to explore analytic techniques which may capture important individual differences within groups.

5.2 Implications

The primary finding of latent phenotypes of acute threat (“fear”) acquisition trajectories that are comparable to those identified by Galatzer-Levy and colleagues (2017) increases confidence that the original finding of heterogeneity in fear acquisition was representative of a

characteristic that may be found in the general population rather than being specific to a sample with higher-than-average levels of psychopathology or a result of type I error. This suggests that phenotypes of fear acquisition may be a relatively robust finding evident in diverse populations. In other words, though fear learning research often emphasizes fear extinction and inhibition, a susceptibility to fear acquisition may be a risk factor for fear and anxiety related psychopathology. For example, abnormalities in amygdala activity during fear acquisition are associated with PTSD and a number of other disorders (Bremner et al., 2005). As mentioned previously, measures of fear acquisition index changes in specific neurobiological circuitry which are evident in both human and rodent models. Additionally, fear responses can be measured separately from DSM categories and symptoms. Thus, this finding suggests that differences in susceptibility to fear acquisition may represent intermediate phenotypes and may be worth of further elaboration in regards to their relationship with the development of clinically significant psychological dysfunction.

Additionally, this observation, combined with a previous finding by Galatzer-Levy and colleagues (2017), may challenge a popular notion of the etiology of PTSD. Emotional Processing Theory (EPT) is built upon the notion that all individuals initially develop dysfunctional cognitions after experiencing a trauma and that one must activate their "fear structure" in order to process a traumatic event and recover (Foa & Kozak, 1991). Emotional Processing Theory provides the theoretical bases for exposure therapy. Though the current study does not assess cognitions, by asserting that one must directly exposure oneself to a "trigger" in order to process a trauma, EPT implies that subcortical processes must be critical in forming and treating this maladaptive fear structure. Furthermore, Prolonged Exposure (PE) therapy, which is one of two specific treatments that are strongly recommended by the 2017 APA treatment

guidelines for PTSD and one of two treatments the Veterans Administartion has adopted for use in its clinics, is a practical application of EPT and is heavily based on the assumption that subcortical fear circuitry is a major driving force in PTSD symptoms (Bufka et al., 2017; Reisman, 2016; Rauch & Foa, 2006). However, contrary to the assumption of homogenous initial responding that is made by EPT, this study and others show signs of heterogeneous initial responses to aversive experiences including both aversive stimuli in lab settings and severe traumatic events (Galatzer-Levy et al., 2013a.; Galatzer-Levy et al., 2013b.). Though this study does not disprove EPT given that it does not measure cognition and only measures exposure to a mildly aversive stimulus, it does add to the existing evidence that EPT's characterization of the traumatic response may not be entirely accurate, particularly in assuming homogeneity in the tendency to acquire aversive associations.

On the other hand, the findings of heterogeneity in acquisition within a minority of participants acquiring a fear while the majority habituates may provide support for Ehlers and Clark's Cognitive Model (Ehlers & Clark, 2000). Though this model is largely focused on cognition, it points to pre-traumatic and peri-traumatic factors as the initial temporal factor in a sequence of events which ultimately may lead to PTSD in some but not others. More specifically, this model notes that relevant peri-traumatic factors are primarily subconscious and experiential in nature, rather than related to conceptual processing (Ehlers & Clark, 2000). Individual differences in susceptibility to fear acquisition may represent a pre-traumatic risk factor and may also indicate a sensitivity to aversive stimuli that may be partially explained by a tendency to process this information differently, making it a potential peri-traumatic factor as well. Therefore, these findings may provide support for the Ehlers and Clark model.

The failure to confirm hypotheses regarding individual differences in fear extinction has several implications. Due to the inability to isolate and test specific probable causes of this discrepancy, firm conclusions cannot be drawn at this time. However, this does demonstrate the importance of interpreting findings from any single GMM study with caution, especially when sample sizes are small. One tangible implication is that future studies will require larger sample sizes in order to examine extinction trajectories. Another potential implication is that, in spite of the common focus on differences in extinction, the finding of individual differences in fear acquisition may be more robust since it was replicated in this sample while differences in extinction were not. Though this may not seem like a likely cause of distress if these individuals can extinguish fear effectively, research shows that fear tends to rapidly generalizes across conceptually related stimuli while extinction is not as readily generalized. Extinction generalization requires successful identification of the initial CS, which is easily tested in lab settings. Though it is easily tested in lab settings, identification of the initial CS may not always be feasible in practical applications such as therapeutic treatment of clients with generalized anxiety that lacks an obvious origin or with multiple early life traumas. Thus, acquired fears are easily generalized and these generalized fears may be difficult to treat (Vervoot et al., 2014). Since this study used a paradigm that was designed to maximize discrimination of CS+ and CS-, fear generalization is unlikely to have occurred. However, participants who easily extinguish fear during discriminant conditioning paradigms may struggle to extinguish fears that are conditioned during lived experience if they have had time to generalize. Thus, an implication of this study is that more attention should be paid to individuals who acquire fear easily as well as specific conditions which may alter extinction (e.g. generalized fears) and the consequences that this may have.

The additional findings derived from ANOVA further demonstrate the need for analytic innovations in biological psychology research. One broad implication of this study's findings is that previous findings utilizing central tendencies statistics to examine fear learning may be missing important individual differences within the groups being compared. Thus, while such studies provide valuable information about group differences, it is important to avoid assuming that results apply to each individual group member. The finding of latent heterogeneity has important implications for clinical practice. Many current therapies and diagnostic techniques are designed based on studies of central tendencies. Thus, heterogeneity in latent characteristics, including basic neurobiological processes, could partially explain the variability in patient response to treatment and development of assessment measures which account for this heterogeneity and treatments which target the associated dysfunctional circuitry may increase the rate of positive responses to therapy. Additionally, since central tendencies analyses of fear learning studies have been applied to a wide array of topics both within and outside of clinical psychology, this means that current knowledge may underestimate individual differences and that GMM approaches may be useful in exploring hitherto unexamined heterogeneity in a wide array of archival data.

This study's identification of latent classes in CS+ but not CS- trials is consistent with findings from Galatzer-Levy and colleagues (2017). This adds to the evidence that normal GMM models may be useful in non-normal data and should not be entirely discarded in response to the advent of the newly developed sktest option for GMM in Mplus. This may have implications for future GMM studies as it provides researchers with more evidence from which to base statistical decision making (e.g. a researcher who must make a difficult decision regarding which distribution to use in a dataset of non-normal physio data may draw upon this stud as evidence to

support the exploration and interpretation of a normal distribution model). Additionally, the finding that this study's relatively mild UCS was effective overall may have implications in further validating the use of a safe and free stimulus which may allow researchers with limited access to more expensive equipment to set up fear learning studies.

5.3 Future Directions

One important future direction for research in this area is for studies to continue to explore and refine the concept of fear learning phenotypes. First, replication studies should be conducted to validate results in the current literature. Ideally, this should be done utilizing pooled archival data to generate large samples and enable retrospective multisite replication studies. In addition to direct replication, conceptual replication across study designs and populations is also needed in order to probe external validity and identify boundary conditions.

Given that multiple studies suggest that a substantial minority of individuals display an enhanced susceptibility to fear, studies should further explore specific differences in biological mechanisms among individuals with the susceptible phenotype. Based on the replicated finding of a latent group of individuals who are unusually susceptible to acquiring fear, rodent and human studies examining the role of glutamate, NMDA receptors, and Basolateral Amygdala activity in shaping individual differences in the likelihood of evidencing a phenotype of increased sensitivity to fear acquisition may prove fruitful. Research should also explore potential target mechanisms and approaches for modulating susceptibility to fear acquisition. The 2017 study by Galatzer-Levy and colleagues demonstrates the potential of GMM applications to rich archival datasets by analyzing pooled data from a variety of human and animal studies in order to link fear extinction profiles in humans with the stress related gene FKBP5 while

separately linking extinction in mice with deficits in FKBP5 mRNA production and further demonstrating that these deficits can be eliminated in rodents with dexamethasone. Drawing inspiration from studies such as that one, future studies should seek ways to make use of archival data to test new hypotheses across several related but separate pooled samples.

Additionally, future studies should examine conditions under which this phenotype may or may not confer increased risk for psychopathology. For example, though the individuals in this study were able to extinguish fear when presented with an ideal opportunity to extinguish fear to the exact same stimulus, paradigms which measure fear generalization may capture phenotypes of fear generalization which are not captured in this study. Mixture modeling and other methods of detecting individual differences should also be extended to other aspects of fear learning including recall, safety signal inhibition, uncertain threat ("anxiety"), and fear generalization. Such studies may have specific implications and applications that are unaccounted for by relatively simple discriminant learning paradigms and may help to generate and refine hypotheses targeting specific mechanisms of fear learning. Paradigms such as the Neutral Predictable Uncertain (NPU) threat task are ideally suited for examining the RDoC construct of uncertain threat ("anxiety") (Schmitz & Grillion, 2012). Ideally, researchers should seek to utilize the current wealth of archival fear learning data in order to increase logistic feasibility while also increasing sample sizes and allowing for single experiments which synthesize across designs.

In addition, the application of GMM to fear learning data in order to detect phenotypes may open the door for a wide array of studies with important implications for a wide range of psychological topics outside of clinical psychology. For example, fear learning studies

examining intergroup relations demonstrate that perception of acute threat ("fear") may be more condition-able to racial, political, and other outgroups as well as less easily extinguished. However, individual differences in this effect have not yet been studied. Thus, the degree to which this effect is universal is unclear. Future studies may be able to apply GMM techniques in order to examine this question and others.

5.4 Limitations

One inherent limitation of a GMM approach is that it is exploratory in nature and thus results must be replicated numerous times before they can reliably form the basis for a given theory. Additionally, this particular study employs a small sample size and utilizes linear modeling and a normal distributional assumption to investigate non-normal data, which may further decrease external validity. Furthermore, the small sample size necessitated the use of models with constrained parameters and extinction trials were dropped for piecewise models due to local solutions in models which included all trials. Thus, important information such as within-class variability and late extinction blocks were excluded from analysis. From an RDoCian perspective, this prevents the analysis from being truly dimensional.

The current study design differs from the only other known human study of GMM during fear learning in a wide variety of ways including the use of LCGA rather than LGMM, sample characteristics, and several key aspects of experimental design. Thus, while it is informative in further supporting the likelihood of distinct fear learning phenotypes, the reason for divergence in extinction findings cannot be isolated. Thus, the finding of one acquiring phenotype which extinguishes fear quickly rather than two which evidence differences in extinction should be interpreted with caution.

An additional limitation of this study is that it does not test the applicability to treatment or the degree to which these phenotypes may relate to clinical difficulties.

5.5 Conclusion

In conclusion, this study adds to the prior literature and represents additional proof of concept for the application of GMM methods to fear learning data in order to examine phenotypes of the RDoC construct of acute threat ("fear"). More specifically, this study suggests that a small minority of individuals may be more susceptible to conditioned fear responses, which may have implications for clinical theory and practice in addition to other topics.

References

Ashbaugh, A. R., Houle-Johnson, S., Herbert, C., El-Hage, W., & Brunet, A. (2016, October 10). Psychometric Validation of the English and French Versions of the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5). *PLoS One, 11(10).* doi: 10.1371/journal.pone.0161645

Asmundson, G. J., Frombach, I., McQuaid, J., Pedrelli, P., Lenox, R., & Stein, M. B. (2000). Dimensionality of posttraumatic stress symptoms: A confirmatory factor analysis of DSM-IV symptom clusters and other symptom models. *Behaviour Research and Therapy*, 38, 203–214.

Beck, A. T. (1990, 1993). BAI, *Beck anxiety inventory: Manual*. San Antonio, TX: Psychological Corporation.

Benke, C., Blumenthal, T.,D., Mode, C., Hamm, A.O., & Pane-Farre, C.A. (2015). Effects of anxiety sensitivity and expectations on the modulation of the startle eyeblink response during a caffeine challenge. *Psychopharmacology*, 232, 3403-3416

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & Boxtel, A. V. (2005, February 09). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, 42(1), 1-15. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1111/j.1469-8986.2005.00271.x

Brewin, C.R., Gregory, J.D., Lipton, M., Burgess, N. (2010). Intrusive images in psychological disorders: characteristics, neural mechanisms, and treatment implications. *Psychological Review* 1(117). 210-232

Briscione, M.A., Jovanovic, T., & Norrholm, S.D. (2014). Conditioned fear associated phenotypes as robust, translational indices of trauma-, stressor-, and anxiety-related behaviors. *Frontiers in Psychiatry* 82(3)

Bufka, L. F., Halfond, R., & Krutzman, H. (2017, February 24). Clinical Practice Guidline for the Treatment of PTSD. American Psychological Association

Carlson, E.B., Smith, S.R., Palmieri, P.A., Dalenberg, C.J., Ruzek, J.I., Kimerling, R., Burling, T.A., & Spain, D.A. (2011). Development and validation of a brief self-report measure of trauma exposure: The Trauma History Screen (PDF). *Psychological Assessment*, *23*, 463-477. doi: 10.1037/a0022294

Centers for Disease Control Vietnam Experience Study Group. (1988). Health status of Vietnam veterans: Psychosocial characteristics. *Journal of the American Medical Association*, 259, 2701–2707.

Christianson, J. P., Fernando, A. B. P., Kazama, A. M., Jovanovic, T., Ostroff, L. E., & Sangha, S. (2012). Inhibition of Fear by Learned Safety Signals: minisymposium review. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(41), 14118–14124. http://doi.org/10.1523/JNEUROSCI.3340-12.2012

Creamer, M., Foran, J., & Bell, R. (1995, May). The Beck Anxiety Inventory in a non-clinical

sample. *Behavior Research and Therapy, 33(4),* 477-485. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7755538

Davis, M. (2006). "Neural systems involved in fear and anxiety measured with fear-potentiated startle." *The American Psychologist, 61(8), 741-756*

Davis M, Walker DL, Myers KM (2003). Role of the amygdala in fear extinction measured with potentiated startle. *Ann N Y Acad Sci.*;985:218–232.

Efekhari, A. Ruzek, J.I. Crowley, J.J. & Rosen C.S. (2013). "Effectiveness of National Implementation of Prolonged Exposure Therapy in Veterans Affairs Care." *JAMA Psychiatry. 70(9):949-955*

Ehlers, A., & Clark, D.M. (2000). "A cognitive model of posttraumatic stress disorder." *Behavioral Research and Therapy.* 38(4); 319-45

Elhai, J., Miller, M., Ford, J., Biehn, T., Palmieri, P., & Frueh, B. (2012). Posttraumatic stress disorder in DSM-5: Estimates of prevalence and symptom structure in a nonclinical sample of college students. *Journal of Anxiety Disorders,* 26(1), 58-64. doi:10.1016/j.janxdis.2011.08.013

Foa, E. B., & Kozak, M. J. (1991). Emotional processing: Theory, research, and clinical implication for anxiety disorders. In J. D. Safran & L. S. Greenberg (Eds.), *Emotion, psychotherapy and changes* New York: Guilford.

Foa, E. B., Kozak, M. J., Salkovskis, P. M., Coles, M. E., & Amir, N. (1998). The validation of a

new obsessive–compulsive disorder scale: The Obsessive–Compulsive Inventory. *Psychological Assessment, 10(3),* 206-214. http://dx.doi.org/10.1037/1040-3590.10.3.206

Foa, E. B., Zinbarg, R., & Rothbaum, B. O. (1992). Uncontrollability and unpredictability in post-traumatic stress disorder: An animal model. *Psychological Bulletin, 112*(2), 218-238. Retrieved from http://dx.doi.org/10.1037/0033-2909.112.2.218

Forster, K.I. & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavioral Research Methods, Instruments, & Computers,* 35(1), 116-124

Galatzer-Levy, I.R., Ankri, Y., Freedman, S., Israeli-Shalev, Y., Roitman, P., Gilad, M., & Shalev, A.Y., (2013). "Early PTSD symptom trajectories: Persistence, recovery, and response to treatment: Results from the jerusalem trauma outreach and prevention study (J-TOPS)." *PLoS One, 8(8) 203-217*

Galatzer-Levy, I. R., & Bryant, R. A. (2013). 636,120 ways to have posttraumatic stress disorder. *Perspectives on Psychological Science,* 8(6), 651-662. doi:10.1177/1745691613504115

Galatzer-Levy, I.R., Bonanno, G.A., Bush, D.E., & LeDoux J.E., (2013). "Heterogeneity in threat extinction learning: substantive and methodological considerations for identifying individual difference in response to stress." *Frontiers in Behavioral Neuroscience, 59(7). 26-28*

Galatzer-Levy, I.R., Andero, R., Sawamura, T., Jovanovic, T., Papini, S., Ressler, K.J., & Norrholm, S.D. (2017). A cross species study of heterogeneity in fear extinction learning

in relation to FKBP5 variation and expression: Implications for the acute treatment of posttraumatic stress disorder, *Neuropharmacology*, 116

Glenn, C. R., Lieberman, L., & Hajcak, G. (2012, September 21). Comparing electric shock and a fearful screaming face as unconditioned stimuli for fear learning. *International Journal of Psychophysiology, 86(3),* 214-219. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0167876012005910

Grady, A. K., Bowen, K. H., Hyde, A. T., Totsch, S. K., & Knight, D. C. (2016, February). Effect of continuous and partial reinforcement on the acquisition and extinction of human conditioned fear. *Behav Neurosci., 130(1),* 36-43. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26692449

Gren-Landell, M., & Universitetet i Linköping. (2010). Social anxiety disorder in swedish adolescents: Prevalence, victimization & development, *Journal of Anxiety Disorders,* 19 (34)

Grillon, C., & Davis, M. (2007, January 30). Fear-potentiated startle conditioning in humans: Explicit and contextual cue conditioning following paired versus unpaired training. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1111/j.1469-8986.1997.tb02389.x

Helzer, J. E., Robins, L. N., & McEvoy, L. (1987). Post-traumatic stress disorder in the general population: Findings of the Epidemiologic Catchment Area Survey. *The New England Journal of Medicine*, 317, 1630–1634. doi:10.1056/nejm 198712243172604

Holtzman-Assif, O., & Laurent, V. (1970, January 01). Orit Holtzman-Assif. *Learning and Memory, 25(4),* 71-75. Retrieved from http://learnmem.cshlp.org/content/17/2/71.short

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167, 748–751. doi:10.1176/appi.ajp.2010.09091379

Izgic, F., Akyuz, G., Dogan, O., & Kugu, N. (2004). Social phobia among university students and its relation to self-esteem and body image. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie*, 49(9), 630-634.

Kim, J., An, B., Kim, J., Park, S., Park, S., Hong, I., Lee, S., Park, K., & Choi, S. (2015). mGluR2/3 in the Lateral Amygdala is Required for Fear Extinction: Cortical Input Synapses onto the Lateral Amygdala as a Target Site of the mGluR2/3 Action. *Neuropsychopharmacology*, 40(13), 2916-28.

Kozak, M. J., & Cuthbert, B. N. (2016). The NIMH research domain criteria initiative: Background, issues, and pragmatics. *Psychophysiology*, 53(3), 286-297. doi:10.111 P.J.

Lang, L.M. McTeague, & M.M. Bradley (2016). "The NIMH research domain criteria initiative: Background, issues, and pragmatics." *Psychophysiology*. 53(3), 286-297 1/psyp.12518

Lanius, R., Bluhm, R., Lanius, U., & Pain, C. (2006). A review of neuroimaging studies in PTSD: Heterogeneity of response to symptom provocation. *Journal of Psychiatric Research*, 40(8), 709-729. doi:10.1016/j.jpyschires.2005.07.007

Lebrón, K., Milad, M. R., & Quirk, G. J. (1970, January 01). Delayed Recall of Fear Extinction in Rats With Lesions of Ventral Medial Prefrontal Cortex. *Learning and Memory, 25(4),* 544-548. Retrieved from http://learnmem.cshlp.org/content/11/5/544.short

Lee, S., Ahmed, T., Lee, S., Kim, H., Choi, S., Kim, D.S., Kim, S.J., Cho, J., & Shin, H.S. (2012). Bidirectional modulation of fear extinction by mediodorsal thalamic firing in mice. *Nature Neuroscience*, 63(9) doi:10.1038/nn.2999

Mahan, A. L., & Ressler, K. J. (2012, January). Fear Conditioning, Synaptic Plasticity, and the Amygdala: Implications for Posttraumatic Stress Disorder. *Trends Neurosci., 35(1),* 24-35. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3206195/

Malik, M. (1996). Standards of Measurement, Physiological Interpretation, and Clinical Use: Task Force of The European Society of Cardiology and the North American Society for Pacing and Electrophysiology. *Annals of Noninvasive Electrocardiology,* 1(2), 151-181

Merz, C. J., Hamacher-Dang, T. C., & Wolf, O. T. (2016, March 16). Immediate extinction promotes the return of fear. *Neurobiology of Learning and Memory, 131*, 109-116. Retrieved from https://www.sciencedirect.com/science/article/pii/S107474271630003X

Mineka, S., & Oehlberg, K. (2008, January 28). The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta Psychologica, 127(3),* 567-580. Retrieved from https://www.sciencedirect.com/science/article/pii/S0001691807001369

Muthen, B., & Asparouhov, T. (2015). Growth mixture modeling with non-normal distributions. *Statistics in Medicine,* 34(6), 1041-58

Muthén, L. K., & Muthén, B. O. (1998-2011). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

National Institute of Mental Health. (n.d.). Retrieved March 1, 2017, from https://www.nimh.nih.gov/research-priorities/rdoc/units/index.shtml

S.D. Norrholm, E.M. Glover, J.S. Stevens, N. Fani, I.R. Galatzer-Levy, B. Bradley, K.J. Ressler, & T. Jovanovic (2015). "Fear Load: The Psychophysiological Over-expression of Fear as an Intermediate Phenotype Associated with Trauma Reactions." *International Journal of Psychophysiology, 98(0),270-275*

Norrholm, S. D., Jovanovic, T., Olin, I. W., Sands, L. A., Karapanou, I., Bradley, B., & Ressler, K. J. (2011). Fear extinction in traumatized civilians with posttraumatic stress disorder: Relation to symptom severity. *Biological Psychiatry,* 69(6), 556-563. doi:10.1016/j.biopsych.2010.09.013

Park, E., Cho, M., & Chang-Seok. (2009, February 28). Correct Use of Repeated Measures Analysis of Variance. *The Korean Journal of Laboratory Medicine, 29(1),* 1-9. Retrieved

from https://synapse.koreamed.org/search.php?where=aview&id=10.3343/kjlm.2009.29.1.1&code=0039KJLM&vmode=AR

Patrick, C.J., & Hajcak, G. (2016). RDoC: Translating Promise into Progress. *Psychophysiology*, 53(3), 415-424 https://doi.org/10.1111/psyp.12612

Phelps EA, Delgado MR, Nearing KI, LeDoux JE. Extinction Learning in Humans: Role of the Amygdala and vmPFC. *Neuron*. 2004;43:897–905.

Powers, M. B., Halpern, J. M., Ferenschak, M. P., Gillihan, S. J., & Foa, E. B. (2010). A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clinical psychology review*, 30(6), 635-641

Prenoveau, J. M., Craske, M. G., Liao, B., & Ornitz, E. M. (2013, January). Human fear conditioning and extinction: Timing is everything . . . or is it?. *Biol Psychol., 92(1)*, 59-68. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3387342/

Quirk GJ, Beer JS (2006). Prefrontal involvement in the regulation of emotion: convergence of rat and human studies. *Current Opinion in Neurobiology*.;16:723–727.

Quirk, G. J., Russo, G. K., Barron, J. L., & Lebron, K. (2000, August 15). The Role of Ventromedial Prefrontal Cortex in the Recovery of Extinguished Fear. *The Journal of Neuroscience, 20(16),* 6225-6231. Retrieved from http://www.jneurosci.org/content/20/16/6225.short

R Core Team (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org

Ram, N., & Grimm, K. J. (2009, October 26). Methods and Measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development, 33(6).* Retrieved from http://journals.sagepub.com/doi/abs/10.1177/0165025409343765

Reisman, M. (2016). PTSD Treatment for Veterans: What's Working, What's New, and What's Next. *Pharmacy and Therapeutics*, *41*(10), 623–634.

Rauch, S., & Foa, E. (2006). Emotional processing theory (EPT) and exposure therapy for PTSD. *Journal of Contemporary Psychotherapy*, *36*(2), 61-65.

Schottenbauer, M. A., Glass, C. R., Arnkoff, D. B., Tendick, V., & Gray, S. H. (2008). Nonresponse and dropout rates in outcome studies on PTSD: Review and methodological considerations. *Psychiatry*, 71(2), 134-168. doi:10.1521/psyc.2008.71.2.134

Schmitz, A., & Grillion, C. (2012). Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). *Nature Protocols*, 7(3)527-32. doi: 10.1038/nprot.2012.001.

Schry, A., Roberson-Nay, R., & White, S. (2012). Measuring social anxiety in college students: A comprehensive evaluation of the psychometric properties of the SPAI-23. *Psychological Assessment*, 24(4), 846-854. doi:10.1037/a0027398

Sehlmeyer, C., Schöning, S., Zwitserlood, P., Pfleiderer, B., Kircher, T., Arolt, V., & Konrad, C. (2009, June 10). Human Fear Conditioning and Extinction in Neuroimaging: A Systematic Review. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0005865

Sherin, J. E., & Nemeroff, C. B. (2011, September). Post-traumatic stress disorder: The neurobiological impact of psychological trauma. *Diologues Clin Neurosci., 13(3),* 263-278. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3182008/

Simonds, L., Thorpe, S., & Elliott, S. (2000). THE OBSESSIVE COMPULSIVE INVENTORY: PSYCHOMETRIC PROPERTIES IN A NONCLINICAL STUDENT SAMPLE. *Behavioural and Cognitive Psychotherapy, 28(2),* 153-159.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009, August 15). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Res., 168(3),* 242-249. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3474329/

Van Ginkel, J. R., & Kroonenberg, P. M. (2014, February 18). Analysis of Variance of Multiply Imputed Data. *Mulitivariate Behavioral Research, 49(1),* 78-91. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/00273171.2013.855890

Vervoort E, Vervliet B, Bennett M, Baeyens F (2014) Generalization of Human Fear Acquisition and Extinction within a Novel Arbitrary Stimulus Category. *PLOS ONE* 9(5): e96569. https://doi.org/10.1371/journal.pone.0096569

Waters AM, Nazarian M, Mineka S, Zinbarg RE, Griffith JW, Naliboff B, Ornitz EM, Craske MG (2014). Context and explicit threat cue modulation of the startle reflex: Preliminary evidence of distinctions between adolescents with principal fear disorders versus distress disorders. *Psychiatry Research*. 2014;217:93–99.

Watts, B. V., Schnurr, P. P., Mayo, L., Young-Xu, Y., Weeks, W. B., & Friedman, M. J. (2013). Meta- analysis of the efficacy of treatments for posttraumatic stress disorder. The Journal of clinical *psychiatry*, 74(6), e541-50.

Waldon, E. G. (2014). DSM-5: Changes and controversies. Music Therapy Perspectives, 32(1), 78-83. doi:10.1093/mtp/miu011

Weathers, F.W., Litz, B.T., Keane, T.M., Palmieri, P.A., Marx, B.P., & Schnurr, P.P. (2013). *The PTSD Checklist for DSM-5 (PCL-5)*. Scale available from the National Center for PTSD at www.ptsd.va.gov.

Wendt, J., Neubert, J., Koenig, J., Thayer, J.F., & Hamm, A.O. (2015). Resting Heart Rate Variability is Associated with Inhibition of Conditioned Fear, *Psychophysiology*, 52(9)

Weston CS. posttraumatic stress disorder: a theoretical model of the hyperarousal subtype. 2014;5

Wickrama, K. A. S., Lee, T. K., O'Neal, C. W., & Lorenz, F. O. (2016). Multivariate applications series. Higher-order growth curves and mixture modeling with Mplus: A practical guide. New York, NY, US: Routledge/Taylor & Francis Group.

Yuzhe, Nakae, K., Ishii, S., & Naoki, H. (2016, September 12). Uncertainty-Dependent Extinction of Fear Memory in an Amygdala-mPFC Neural Circuit Model. Retrieved from http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005099

Zimmerman, J. M., & Maren, S. (2010, April 23). NMDA receptor antagonism in the basolateral but not central amygdala blocks the extinction of Pavlovian fear conditioning in rats. *European Journal of Neuroscience, 31(9),* 1664-16670. Retrieved from https://onlinelibrary.wiley.com/doi/full/10.1111/j.1460-9568.2010.07223.x

Zoellner, L. A., Pruitt, L. D., Farach, F. J., & Jun, J. J. (2014). understanding heterogeneity in ptsd: Fear, dysphoria, and distress. *Depression and Anxiety*, 31(2), 97-106. doi:10.1002/da.22133

Zoladz, P.R. & Diamond, D.M. (2017). Current Status on Behavioral and Biological Markers of PTSD: A Search for Clarity in a Conflicting Literature. *Neuroscience & Biobehavioral Reviews*, 37(5), 860-895

Appendix

Figure 1. Participant Characteristics

	Sample (n=67)	
	n / M	*% / SD*
Gender		
Female	42	64.4
Male	25	35.6
Race/Ethnicity		
Caucasian	43	66.2
African-American	3	4.6
Hispanic	5	7.7
Asian	3	4.6
Did not provide race	13	16.9
Scores on Self-Report Measures		
THS	3.04	3.76
PCL-5	15.94	15
OCD-I	7.48	4.9
BAI	10.36	9.37
Substance use on Day of Experiment		
Caffeine (within 12 hours)		
Day 1	15	22.4
Day 2	11	23.4
Nicotine (within 12 hours)		
Day 1	2	3.0
Day 2	3	6.4
Alcohol (within 24 hours)		
Day 1	0	0
Day 2	1	2.1

Figure 2. Paradigm

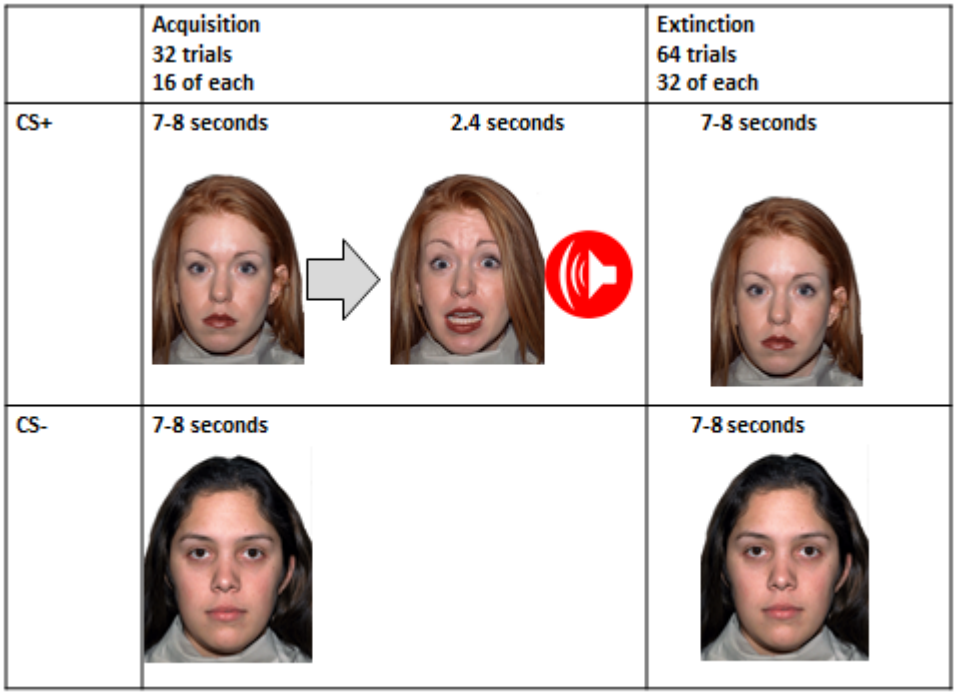

Figure 3. Trial Structure and ITI Structure

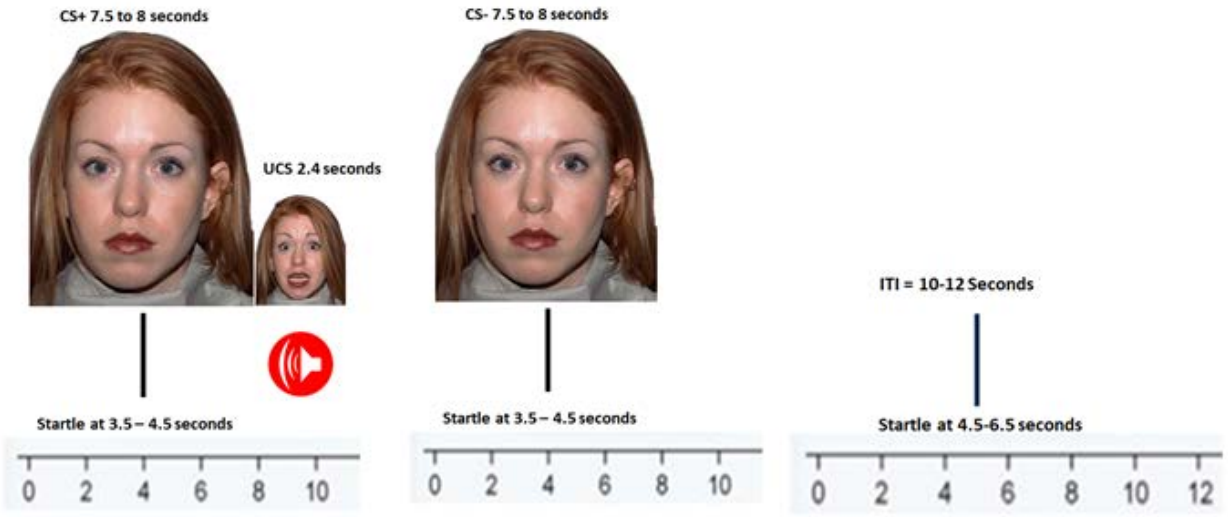

Figure 4. ANOVA Day 1

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Block	Wilks' Lambda	.554	3.882[b]	11.000	53.000	.000	.446
Block * DMDX_mean	Wilks' Lambda	.552	1.063	33.000	156.852	.388	.180
CS	Wilks' Lambda	.573	46.865[b]	1.000	63.000	.000	.427
CS * DMDX_mean	Wilks' Lambda	.966	.732[b]	3.000	63.000	.537	.034
Block * CS	Wilks' Lambda	.696	2.102[b]	11.000	53.000	.036	.304
Block * CS * DMDX_mean	Wilks' Lambda	.495	1.281	33.000	156.852	.160	.209

a. Design: Intercept + DMDX_mean

Within Subjects Design: Block + CS + Block * CS

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Figure 5. ANOVA Day 2

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Block	Wilks' Lambda	.874	.780b	7.000	38.000	.608	.126
Block * DMDX_mean	Wilks' Lambda	.746	.560	21.000	109.666	.937	.093
CS	Wilks' Lambda	.820	9.678b	1.000	44.000	.003	.180
CS * DMDX_mean	Wilks' Lambda	.655	7.738b	3.000	44.000	.000	.345
Block * CS	Wilks' Lambda	.733	1.973b	7.000	38.000	.085	.267
Block * CS * DMDX_mean	Wilks' Lambda	.383	2.070	21.000	109.666	.008	.274

a. Design: Intercept + DMDX_mean

Within Subjects Design: Block + CS + Block * CS

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Figure 6. ANOVA Two Days

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Phase	Wilks' Lambda	.660	5.281b	4.000	41.000	.002	.340
Phase * DMDX_mean	Wilks' Lambda	.829	.664	12.000	108.767	.782	.060
CS	Wilks' Lambda	.797	11.184b	1.000	44.000	.002	.203
CS * DMDX_mean	Wilks' Lambda	.788	3.949b	3.000	44.000	.014	.212
Phase * CS	Wilks' Lambda	.540	8.723b	4.000	41.000	.000	.460
Phase * CS * DMDX_mean	Wilks' Lambda	.594	1.975	12.000	108.767	.033	.160

a. Design: Intercept + DMDX_mean

Within Subjects Design: Block + CS + Block * CS

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Figure 7. Fit Indices Acquisition

Figure 7. Fit Indices for One-to Four Class linear only LCGA Models based on Fear-Potentiated Startle Response Difference Scores to CS + Condition during acquisition (n = 67); 1- to 3-class linear and quadratic LCGA models to CS+ during acquisition (N = 67).

Fit Indices	AIC	BIC	SSBIC	LRT	BLRT	Entropy
			***Acquisition:* Linear**			
1 Class	214.13	227.27	208.38	–	–	–
2 Class	**190.59**	**210.30**	**181.96**	***p* =0.058**	***p* < 0.001**	**0.91**
3 Class	179.92	206.20	168.42	*p* =.27	*p* < 0.001	0.94
			***Acquisition:* Linear + Quadratic**			
1 Class	216.13	231.46	209.42	–	–	–
2 Class	**193.16**	**217.29**	**182.62**	***p* =.25**	***p* < 0.001**	**0.92**
3 Class	185.66	218.50	171.28	*P*=.30	*p* < 0.001	0.97

Note. Best fitting model is highlighted in **Bold**. P-values reflect comparison of k class model with k - 1 class model with the Lo-Mendel-Rubin likelihood ratio test (LRT) and the Bootstrap Likelihood Ratio Test (BLRT). AIC = Akaike information criterion; BIC = Bayesian information criterion; SSBIC = sample size adjusted Bayesian information criterion.

Figure 8. Fit Indices for Piecewise Acquisition and Extinction

Figure 8. Fit Indices for One-to Four Class linear only Piecewise LCGA Models based on Fear-Potentiated Startle Response Difference Scores to CS + Condition during acquisition and extinction (n = 67); 1- to 4-class linear and quadratic LCGA models to CS+ during acquisition and extinction (N = 67).

Fit Indices	AIC	BIC	SSBIC	LRT	BLRT	Entropy
			***Acquisition and Extinction:* Linear**			
1 Class	398.01	421.76	387.14	–	–	–
2 Class	**360.00**	**392.38**	**345.17**	***p* =0.17**	***p* < 0.001**	**0.92**
3 Class	358.93	399.94	340.15	*p* =.29	*p* < 0.001	0.94
			***Acquisition and Extinction:* Linear + Quadratic**			
1 Class	401.20	429.27	388.35	–	–	–
2 Class	**361.01**	**402.03**	**342.23**	***p* =.25**	***p* < 0.001**	**0.96**
3 Class	332.47	386.44	307.76	*P*=.30	*p* < 0.001	0.96

Note. Best fitting model is highlighted in **Bold**. P-values reflect comparison of k class model with k - 1 class model with the Lo-Mendel-Rubin likelihood ratio test (LRT) and the Bootstrap Likelihood Ratio Test (BLRT). AIC = Akaike information criterion; BIC = Bayesian information criterion; SSBIC = sample size adjusted Bayesian information criterion.

Figure 9. Parameter Estimates for Selected Models

	Intercept			Slope 1			Slope 2		
Acquisition Linear	***EST (EST1)***	**SE**	***p***	***EST***	**SE**	***p***	***EST***	**SE**	***p***
Acquiring	0.81 (0.282)	0.15 (0.07)	<.001 (<.001)	0.18	0.06	0.002	-	-	-
Habituating	0.06 (0.242)	0.04 (0.05)	0.092 (<.001)	-0.61	0.02	<.001	-	-	-
Acquisition and Extinction Linear									
Extinguishing	0.68 (0.32)	0.05 (0.06)	<.001 (<.001)	0.12	0.05	<.001	-0.06	0.02	0.004
Habituating	0.035 (0.233)	0.02 (0.04)	0.06 (<.001)	-0.07	0.02	.001	-0.00	0.01	0.971

Note. Est= Estimate; *SE*= Standard Error of the Estimate; Estimated parameters for each latent class include the *intercept* representing measured scores on the y-axis at the last phase of acquisition *(EST1)* indicating the intercept representing measured scores on the y-axis at the first phase of acquisition; *slope 1* representing the degree of linear change across trials during acquisition; *slope 2* representing the degree of linear change across trials during extinction; For tests of acquisition only,

Figure 10. Acquisition Model

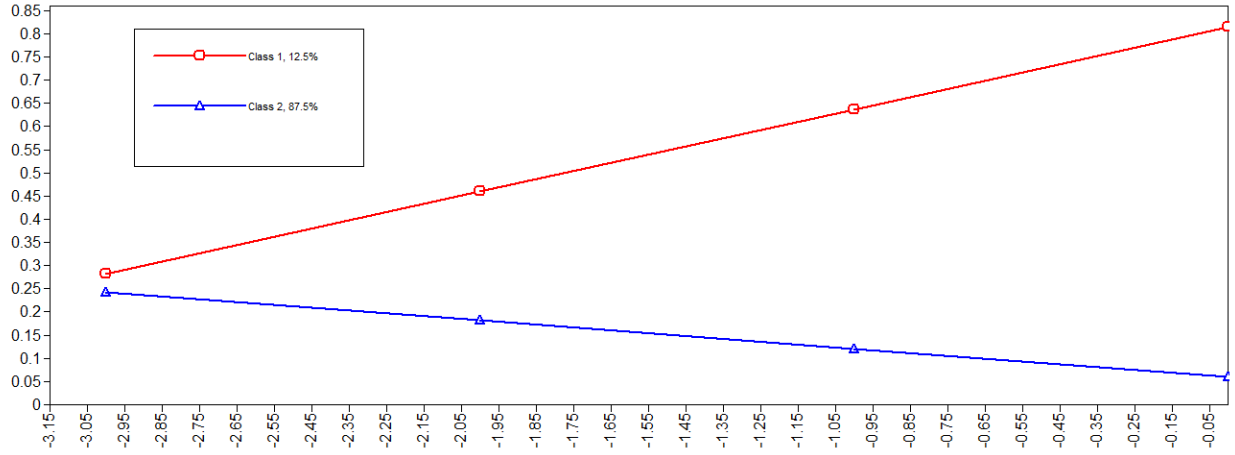

Figure 11. Piecewise Acquisition and Extinction Model