McNair Scholars Research Journal

Volume 14 Article 8

2021

Inhibitory Learning Theory and Emotional Processing Theory for Proper Extinction Learning in Specific Phobias: A Comprehensive Comparison

Arturo Peña
apena1@emich.edu

Inhibitory Learning Theory and Emotional Processing Theory for Proper Extinction Learning in Specific Phobias: A Comprehensive Comparison

Arturo F. Peña
Dr. Ellen Koch, Mentor

ABSTRACT

Emotional processing theory (EPT) is the underlying framework for traditional exposure therapy (Barlow, 2004). EPT relies on a process known as habituation, presuming that the amount of fear reduction from peak fear levels within an exposure therapy session predicts successful extinction learning. Prevailing research in behavioral psychology challenges the EPT framework for its capacity to accurately predict long-term extinction learning. Inhibitory learning theory (ILT) offers a different approach to exposure therapy, centered around forming new non-threat associations (e.g., memories) by way of fear acceptance rather than attempts to down-regulate aversive emotions (Craske et al., 2008). By examining pertinent literature on memory and learning, this review aims to provide insight on the efficacy of the traditional habituation method in comparison to the fear tolerance or fear acceptance methods used in exposure therapy. This paper hopes to add to the growing body of research advancing exposure methodology towards a more unified and successful framework.

INTRODUCTION

The primary purpose of this paper is to examine the efficacy of emotional processing theory (EPT) and inhibitory learning theory (ILT) when used in exposure therapy to address phobias. This paper reviews the history and development of each theory and compares EPT, which relies on fear reduction, to ILT, which relies on fear acceptance. A secondary goal of this paper is to understand the application of EPT and ILT to specific phobias by reviewing studies that test the traditional habituation model and fear acceptance/toleration models.

Specific Phobias: Prevalence and Significance

Specific phobia refers to an intense and persistent fear of a particular object or situation that interferes with one's life (American Psychiatric Association [APA], 2013) and has been associated with aversive outcomes. Common specific phobias include: arachnophobia (fear of spiders), glossophobia (fear of public speaking), acrophobia (fear of heights), and more. Approximately 9.1% of U.S. adults suffer from a specific phobia (National Institute of Mental Health [NIMH], 2017). According to the National Institute of Mental Health, "an estimated 12.5% of U.S. adults experience specific phobia at some point in their lives" (2017, para. 3). A significantly higher percentage of individuals also have intense fears, but they may not meet diagnostic criteria of the fear interfering with their life.

Learning and Pavlovian Conditioning

Researchers began to theorize about the underlying processes that are responsible for human fear learning, and learning in general, during the late nineteenth century. Named for Russian physiologist Ivan Pavlov, Pavlovian conditioning has been one of the primary behavioral frameworks in psychology used to study human learning (Pavlov, 1941). In his dog experiments, Pavlov noted that the animal salivated when presented with food. The food was termed the *unconditioned stimulus,* or UCS. The salivation response was termed as the *unconditioned response,* or UCR. Pavlov then paired a *neutral stimulus* (NS) with the UCS. To do so, he paired a bell (NS) with the delivery of food to the dog (UCS). The NS is neutral because it elicits no response before conditioning occurs, hence its neutrality.

After repeated pairings of the bell with food delivery, the dog began to salivate upon hearing the bell, even before receiving the food. In this stage of Pavlovian conditioning, the bell transitioned from NS to CS because the dog has been conditioned to respond to the bell. The UCR underwent a similar process. A *conditioned response* (CR) occurred after enough repeated pairings of the bell with food delivery such that the bell became the conditioned stimulus. Other important facets of Pavlovian conditioning include *extinction* and *stimulus generalization*. *Extinction* refers to the process by which presentations of the CS without the pairing of the UCS lead to decreased conditional responding. That is, the conditioned stimulus diminishes in its capacity to elicit a conditioned response. If, after conditioning, Pavlov's dog did not receive food after hearing the bell, over time the bell's ability to elicit salivation would de-

crease. *Stimulus generalization* refers to a process by which stimuli that are similar to the conditioned stimulus will elicit responses similar to that of the conditioned stimulus. Following from the Pavlovian example, *stimulus generalization* would occur if the dog salivated due to a tone similar to that of the CS.

Pavlov's research provided many of the foundations of theories of learning and conditioning. Joseph Wolpe was a South African psychiatrist who was influenced both by Ivan Pavlov and John B. Watson, the psychologist who originated the behaviorist model of psychology. Wolpe practiced in the mid-twentieth century and took great interest in the Neo-behaviorist tradition of the period (Craske et al., 2006). Primarily interested in the principles of learning, his research with cats provides insight into the development of systematic desensitization.

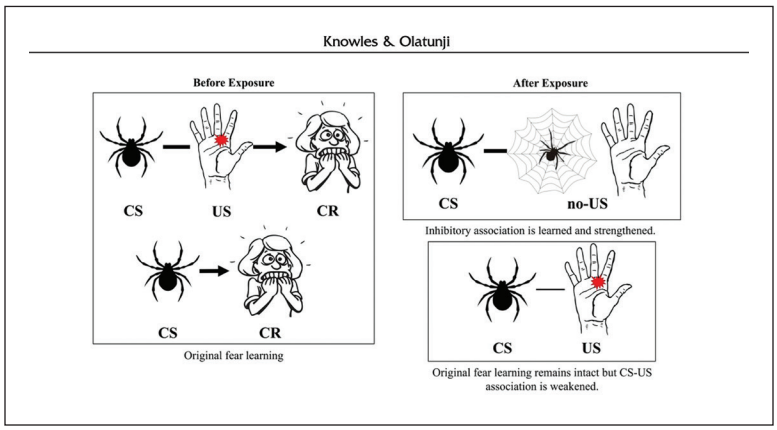
> These animals were given mild electric shocks accompanied by specific sounds and visual stimuli. Once the cats knew to equate the unpleasant shock with these images or sounds, the images and sounds created a feeling of fear. By gradually exposing the cats to these same sights and sounds-with food being given instead of shocks—the cats gradually "unlearned" their fear. (Berger, 2005, para. 4)

This experiment helped form the basis of systematic desensitization along with the work of Pavlov and Watson.

Fear Learning: Acquisition of Specific Phobias

The need for researchers to study phobias is evident, given the number of individuals who experience debilitating fear at some point in their lives. Playing a critical role in human development (Lebois et al., 2019), fear is an adaptive response that has helped us survive throughout our evolutionary history by alerting us to potential dangers. In this sense, fear is functional to survival. A fear response is appropriate when a threat is present because it triggers an avoidance response to the threatening object or situation. However, fear responses can become problematic when no threat is present. Following the Pavlovian paradigm, a CS is paired with a UCS, which leads to a CR. In specific phobias, the CS represents the feared object and the UCS represents the feared outcome. The CR is the fear response.

Figure 1. *Fear associations in the inhibitory learning paradigm (Knowles & Olatunji, 2019).*



In Figure 1, the spider represents the feared object (Knowles and Olatunji, 2019). The hand that has the red aversive symbol on the palm represents the feared outcome. A fear association is formed by the pairing of the feared object and feared outcome, eliciting a fear response. Arachnophobia occurs when the spider, the conditioned stimulus, elicits a fear response without the unconditioned stimulus, the bite, occurring. This type of learning can pathologize and become extremely problematic for individuals who experience this kind of fear learning. For over 40 years, researchers have studied the basis of fear learning in order to design an effective treatment both to specific phobias and to other anxiety disorders that interfere with normal life.

LITERATURE REVIEW

History and Development of Emotional Processing Theory

EPT has been the traditional framework for exposure therapy in the treatment of specific phobias (Craske, 2015; Foa & Kozak, 1986). The emotional processing model was constructed to explain how fear responses diminish and how fear associations are weakened (Foa & Kozak, 1986). Two primary methods of the emotional processing framework have been developed to explain the reduction of fear through the incorporation of incompatible information. These methods include within-session habituation as an accurate predictor of outcome, and between-session habituation as a necessary extension of within-session habituation.

EPT and Fear Structures

EPT theory asserts that human fear structures must be modified. According to Foa and Kozak (1986), the fear structure consists of cognitive representations of the original fear context, the individual's associated responses, and the meaning of the fear association. Baker et al. (2010), citing Lang's (1971) earlier work, described a fear structure as "a set of propositions about a stimulus (e.g., spider), response (e.g., racing heart) and their meaning (e.g., "I will be poisoned") that are stored in memory" (p. 1). When activated, the fear structure elicits fearful responding. EPT aims to activate the fear structure and progressively encode incompatible information to modify the fear associations in a hierarchical fashion, hence the concept of fear reduction, which is synonymous with habituation in this context.

Within-session habituation. *Habituation* refers to the gradual reduction of fear and anxiety levels from initial fear activation (IFA). In exposure, participants are presented with the stimulus they fear to deliberately activate a fear response. Different measures, such as the subjective units of distress scale (SUDS), can be used to subjectively assess the point when the fear structure has been activated. SUDS measures subjective fear from one moment to the next. Within-session habituation (WSH) refers to the amount to which fear decreases from peak fear levels within one exposure session (Foa & Kozak, 1986). According to the EPT paradigm, WSH is a necessary prerequisite for between-session habituation, which is hypothesized to lead to long-term extinction learning.

Between-session Habituation. *Between-session habituation* (BSH) refers to the level of fear reduction between exposure sessions (Foa & Kozak, 1986). In EPT, WSH and BSH are inextricably linked. The theory asserts that fear reduction levels measured within and between sessions is an accurate predictor of successful extinction learning (Thompson, 2009).

Extinction Learning

In Pavlovian classical conditioning, extinction occurs when the CS is no longer paired with the aversive stimulus (Lebois et al., 2019). Figure 1 shows that, after exposure, the pairing of the CS and US lessens to the point of extinction. The fear associated with the spider will gradually diminish as non-threatening interactions occur. Inhibitory learning theory (ILT), by contrast, offers an alternative approach to extinction learning.

History and Development of Inhibitory Learning Theory

The development of ILT is directly related to the history and development of EPT. Similar to EPT, ILT extends from the Pavlovian classical conditioning paradigm (Bouton, 1993; Craske, 2015). ILT seeks the formation of new non-fear associations to promote extinction learning. Within the ILT conceptual framework, it is understood that the original CS-US fear association is not erased from memory. Rather, the original fear association remains intact, but new, non-fear associations will outcompete the original fear association, thereby diminishing the phobia. ILT seeks to inhibit the original fear association (i.e., the CS-US association) by encoding the new non-fear association (i.e., the CS-noUS association).

Figure 1 illustrates the formation of a new, non-fear association. The drawing of the spider without the hand being bitten represents the CS-noUS association. Here, inhibitory learning is occurring. The original fear association remains in memory, however, the new non-fear association lessens the original CS-US fear association. Forming new competing associations instead of focusing on fear reduction (i.e., habituation) distinguishes ILT from EPT. This distinction requires an understanding of fear toleration or fear acceptance as central processes of ILT. Within EPT, fear reduction relies on the gradual decline of fear levels for the original fear association's modification. However, the ILT paradigm relies on fear acceptance to occur so that the new non-fear association can be encoded into the memory of the phobic individual. This toleration of fear is critical for proper extinction learning to occur.

Efficacy of Emotional Processing Theory

A study by Baker et al. (2010) examined the validity of the primary propositions of EPT in acrophobic individuals. These propositions included within-session habituation as an accurate predictor of outcome, and between-sessions habituation as a necessary extension of within-session habituation. Specifically, they studied the value of combining habituation and extinction learning to better understand the effects of exposure therapy. The sample consisted of 44 participants (32 female and 12 male) with a mean age of 18.88 years. Data were collected on four separate instances: baseline, first exposure, second exposure, and two-week post-assessment. Behavioral approach tests (BATs) were conducted at baseline, immediately after the final exposure session, and two-week post-assessment. Throughout the BATs, SUDS were rated at minute 0 and each minute thereafter. Heart rate data were collected throughout

the BATs. This study concluded that the premise of EPT failed to produce significant results to support the accuracy of the theory. No evidence was found for WSH as a predictor of outcomes, and no relationship between WSH and BSH was observed. The authors point out that WSH not being predictive of outcomes or BSH is especially problematic for EPT because within the framework of EPT, within-session habituation is a necessary prerequisite for between-sessions habituation (Baker et al., 2010).

Another group of researchers conducted a meta-analysis of the association between process and outcome measures in emotional-processing-based exposure therapy (Rupp et al., 2017). The researchers aimed to examine statistical correlations of the outcome of exposure therapy and three variables of emotional processing theory, including initial fear activation, within-session habituation, and between-session habituation. This meta-analysis had criteria including studies in which the populations were diagnosed with anxiety disorders; exposure was their primary treatment (including in vivo exposure, imaginal exposure, and VR exposure), and at least one of the three variables' processes (i.e., IFA, WSH, BSH) of the emotional processing framework was statistically correlated with at least one measure assessing outcome. Furthermore, the sample size of any studies included in the meta-analysis had to be at least five *(N=5)*. In the end, 21 studies were included in the meta-analysis. The authors found no statistically significant relationships between the predictive variables of emotional processing theory and outcomes after exposure therapy. Specifically, initial fear activation was not shown to be significantly correlated with outcome measures, challenging the efficacy of emotional processing, which relies on the activation of the fear structure in order for emotional processing to occur.

The EPT framework requires a strict adherence to WSH following initial fear activation, followed by BSH (IFA → WSH → BSH). This strict aspect of the succession of factors is EPT's conceptual limitation. Specifically, there is little evidence suggesting that the amount to which fear diminishes from peak levels is predictive of successful extinction learning (Craske et al., 2008). In other words, fear reduction levels do not have predictive strength. ILT appears to offer more accurate methods for promoting proper extinction learning because ILT does not rely on fear reduction or habituation, instead relying on fear acceptance.

Efficacy of Inhibitory Learning Theory

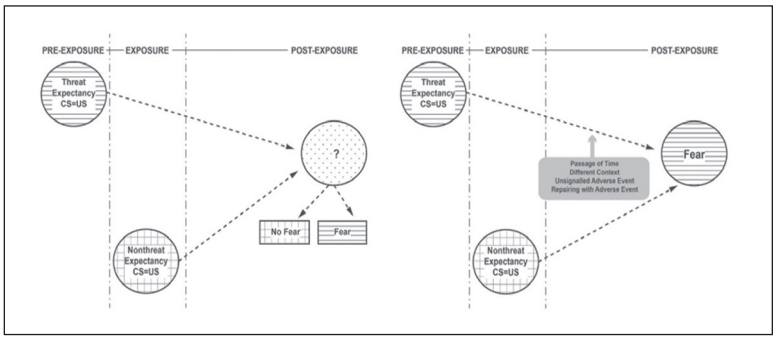
Norberg et al. (2018) designed an experiment to test a central aspect of ILT. The researchers wanted to examine if an extra, fear-inducing

step added to the exposure process would lead to greater distress tolerance in arachnophobic individuals. Their central hypothesis was that a more challenging exposure would result in greater changes in fear and fear beliefs as compared to a less intensive exposure. There were two conditions: the same-context group underwent assessment and treatment in a living room, and the different-context group experienced assessment in a living room and underwent treatment in a therapeutic office. Participants were then categorized into one of two groups. Those who were only able to complete behavioral approach task steps up to 14 or lower were a part of the "step 14 or lower" group. Those individuals who were able to complete the additional step were a part of the "step 15" group, which was the group that experienced a more difficult exposure process.

Participants completed a behavioral approach task before their first exposure session and after their second exposure session. Researchers found support for their initial hypothesis, postulating that a more difficult exposure session would be more effective in promoting extinction learning than a less intensive exposure session. Both treatment groups experienced substantial decreases in their self-reported fear. However, those who completed the extra step reported that their most feared scenarios were less likely to occur. Additionally, after a change in context, these individuals did not lose their capacity to tolerate their worst-case scenarios, unlike those individuals who did not complete the extra step. The central tenet of the ILT paradigm was supported by the data, which showed that toleration of fear, which was required to complete step 15 in the study, was necessary for proper extinction learning to occur. The participants in the step 15 group were able to encode a new non-fear association.

Figure 2 shows the clinical translation of ILT. Both the right and left diagrams of Figure 2 are split into the following three sections: pre-exposure, exposure, and post-exposure. "Pre-exposure" represents a threat expectancy, which is the fear association (i.e., the CS-US association). During "exposure" the individual forms the non-threat expectancy (i.e., the CS-noUS association). "Post-Exposure" shows that either the non-fear or fear association will be retrieved by the phobic individual. Hence the importance of deeply encoding the non-fear association within the ILT framework. The diagram on the right side of Figure 2 shows factors that may lead to the fear association being retrieved over the non-fear association, including passage of time, different contexts, unexpected adverse events, and repairing with an adverse event.

Figure 2. *Craske's clinical translation of the inhibitory learning theory extinction paradigm (2015).*



Repeated encounters with the feared stimuli without aversive events occurring is central to ILT, and not only promotes new learning, but also behavioral change or decreased conditional response (Jacoby & Abramowitz, 2016). This is because the original fear association loses its strength to elicit fear and anxiety due to the meaning acquired by the CS. After extinction the CS has two meanings: the original CS-US fear memory and the new CS-noUS inhibitory memory (Craske et al., 2012).

DISCUSSION

For decades, emotional processing was the primary model by which exposure therapy was designed. Its reliance on habituation led researchers to pursue fear reduction levels as indicators of successful extinction learning, believing that the original fear association could be slowly manipulated and modified through the incorporation of information that is incompatible with the fear beliefs. EPT required a strict succession of treatment steps: first, the fear structure had to be activated (i.e., IFA), then WSH had to occur in order to trigger BSH, which was presumed to lead to proper extinction. The development of the inhibitory learning model showed that a new non-fear memory could be encoded into memory, replacing the earlier fear memory. Additionally, ILT moves away from the habituation approach and focuses on fear acceptance. There is scant literature suggesting that fear reduction levels accurately predict extinction outcomes. However, fear acceptance allows participants to tolerate the feared stimuli so that a new CS-noUS association can be formed. Altogether, ILT is an exciting line of scientific inquiry, but more research is needed to increase the efficacy of the newer

exposure model because there are deficits in extinction learning in those with anxiety disorders (Craske, 2015).

REFERENCES

American Psychiatric Association [APA]. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5(TM))* (5th ed.). American Psychiatric Publishing.

Baker, A., Mystkowski, J., Culver, N., Yi, R., Mortazavi, A., & Craske, M. G. (2010). Does habituation matter? Emotional processing theory and exposure therapy for acrophobia. *Behaviour Research and Therapy,* 48(11), 1139–1143. https://doi.org/10.1016/j.brat.2010.07.009

Barlow, D. H. (2004). *Anxiety and Its Disorders, Second Edition: The Nature and Treatment of Anxiety and Panic* (Second ed.). The Guilford Press.

Berger, V. F. (2005). *Psychologists: Joseph Wolpe | PsychologistAnywhereAnytime.com.* Psychologistanywhereanytime. https://www.psychologistanywhereanytime.com/famous_psychologist_and_psychologists/psychologist_famous_joseph_wolpe.htm

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114*(1), 80–99. https://doi.org/10.1037/0033-2909.114.1.80

Craske, M. G. (2015). Optimizing exposure therapy for anxiety disorders: An inhibitory learning and inhibitory regulation approach. *Verhaltenstherapie, 25*(2), 134–143. https://doi.org/10.1159/000381574

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy, 46*(1), 5–27. https://doi.org/10.1016/j.brat.2007.10.003

Craske, M. G., Liao, B., Brown, L., & Vervliet, B. (2012). Role of Inhibition in Exposure Therapy. *Journal of Experimental Psychopathology, 3*(3), 322–345. https://doi.org/10.5127/jep.026511

Craske, M. G., Hermans, D., & Vansteenwegen, D. (Eds.). (2006). *Fear and learning: From basic processes to clinical implications.* American Psychological Association. https://doi.org/10.1037/11474-000

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin, 99*(1), 20–35. https://doi.org/10.1037/0033-2909.99.1.20

Jacoby, R. J., & Abramowitz, J. S. (2016). Inhibitory learning approaches to exposure therapy: A critical review and translation to obsessive-compulsive disorder. *Clinical Psychology Review, 49*, 28–40. https://doi.org/10.1016/j.cpr.2016.07.001

Knowles, K. A., & Olatunji, B. O. (2019). Enhancing inhibitory learning: The utility of variability in exposure. *Cognitive and Behavioral Practice, 26*(1), 186–200. https://doi.org/10.1016/j.cbpra.2017.12.001

Lang, P. J. (1971). The application of psychophysiological methods to the study of psychotherapy and behavior modification. In A. Bergin, & S. Garfield (Eds.), Handbook of psychotherapy and behavior change. New York: Wiley

Lebois, L. A. M., Seligowski, A. V., Wolff, J. D., Hill, S. B., & Ressler, K. J. (2019). Augmentation of extinction and inhibitory learning in anxiety and trauma-related disorders. *Annual Review of Clinical Psychology, 15*(1), 257–284. https://doi.org/10.1146/annurev-clinpsy-050718-095634

National Institute of Mental Health [NIMH]. (2017, November 1). *Specific phobia.* https://www.nimh.nih.gov/health/statistics/specific-phobia

Norberg, M. M., Newins, A. R., Jiang, Y., Xu, J., Forcadell, E., Alberich, C., & Deacon, B. J. (2018). The scarier the better: Maximizing exposure therapy outcomes for spider fear. *Behavioural and Cognitive Psychotherapy, 46*(6), 754–760. https://doi.org/10.1017/s1352465818000437

Pavlov, I. P., & Gantt, W. H. (1941). *Conditioned reflexes and psychiatry* (Vol. 2). New York: International publishers.

Rupp, C., Doebler, P., Ehring, T., & Vossbeck-Elsebusch, A. N. (2017). Emotional processing theory put to test: A meta-analysis on the association between process and outcome measures in exposure therapy. *Clinical Psychology & Psychotherapy, 24*(3), 697–711. https://doi.org/10.1002/cpp.2039

Thompson, R. F. (2009). Habituation: A history. *Neurobiology of Learning and Memory, 92*(2), 127–134. https://doi.org/10.1016/j.nlm.2008.07.011