NOTE TO USERS

Page(s) not included in the original manuscript and are unavailable from the author or university. The manuscript was scanned as received.

103-107

This reproduction is the best copy available.

UMI®

TESTING A COGNITIVE MODEL OF IMPLICIT SELF-ESTEEM THROUGH EVALUATIVE CONDITIONING

JODENE ROBIN BACCUS

DEPARTMENT OF PSYCHOLOGY

MCGILL UNIVERSITY, MONTREAL, QUEBEC, CANADA

June 2005

A thesis submitted to McGill University

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Your file Votre référence
ISBN: 978-0-494-21616-3
Our file Notre référence
ISBN: 978-0-494-21616-3

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the help of numerous people. First, I would like to thank my advisor Mark Baldwin, not only for his wonderful guidance but also for providing me with the skills and confidence needed to achieve my goals. It is impossible to quantify the amount of knowledge that I have gained from working with him. Thank you also to the following incredible social psychology faculty at McGill University for their assistance and feedback, and for creating such a positive work environment—Richard Koestner, John Lydon, and Don Taylor. Thank you especially to Richard Koestner for comments on an earlier draft of this dissertation. For their invaluable assistance with data collection, entry, and analysis I would like to thank Christopher Bryan, Eleshia Morrison, and Genevieve Taylor. Special acknowledgement goes to Dominic Packer who programmed the original computer-based conditioning task and who collected the data presented in Study 2 as part of his honours research project at McGill University. My graduate school experience would not have been the same without the encouragement and friendship of my lab mates Stephane Dandeneau and Maya Sakellaropoulo. Thank you also to Stephane for translating the abstract into French. Finally, thank you to David Paul for comments on drafts of this dissertation, and for his unwavering support from beginning to end.

ABSTRACT

Implicit self-esteem is the automatic and unconscious component of self-esteem, which is generally not correlated with more traditional measures of explicit self-esteem. The goal of the research presented in this dissertation was to test a cognitive model of implicit self-esteem. Drawing from interpersonal theories of self-esteem and from theories of evaluative learning, I hypothesized that implicit self-esteem is developed through repeated exposure to pairings of the self with interpersonal rejection or acceptance, leading to unconscious and automatic self--*rejected* or self--*accepted* cognitive associations. I tested this theory using a computer-based conditioning task designed to enhance implicit self-esteem through the repeated pairing of self-relevant information (e.g., first name, birthday) with interpersonal acceptance (i.e., photographs of smiling faces). Support for this hypothesis was found in four studies. Overall, participants who completed the computer-based conditioning task generally had higher scores on the self-esteem Implicit Associations Test compared to participants in a control condition. Furthermore, the conditioning task had no effect on explicit self-esteem scores, providing support for the distinct nature of implicit self-esteem relative to explicit self-esteem. Finally, exploratory analyses showed that contingency awareness was not needed for evaluative conditioning to occur. These findings support the proposed cognitive model of implicit self-esteem and provide a novel computer-based conditioning task for enhancing implicit self-esteem.

RÉSUMÉ

L'estime de soi implicite est une composante automatique et inconsciente de l'estime de soi qui n'est généralement pas corrélative avec les mesures plus traditionnelles d'estime de soi. Le but de la recherche présentée dans cette dissertation est de tester un modèle cognitif de l'estime de soi implicite. Retenant des théories interpersonnelles de l'estime de soi et des théories d'apprentissage évaluatif, je propose que l'estime de soi implicite se développe par l'entremise d'accouplages répétitifs entre le soi et le rejet ou l'acceptation sociale qui, par la suite, développent des associations cognitives automatiques et inconscientes du genre soi--*rejeté* ou soi--*accepté*. J'ai testé cette théorie avec une tâche de conditionnement informatisée qui a comme but d'accroître l'estime de soi implicite par l'entremise d'accoupler répétitivement des informations se reportant au soi (e.g. prénom, date de naissance) à l'acceptation sociale, c'est-à-dire des photos de personnes qui sourient. Quatre études supportent mon hypothèse. Globalement, les participants qui ont complété la tâche de conditionnement avaient, en général, des résultats plus élevés sur le test d'associations implicite (IAT) d'estime de soi comparer aux participants qui ont complété la condition contrôle. De plus, le conditionnement n'a pas eu d'effet sur l'estime de soi explicite, appuyant ainsi le fait que l'estime de soi implicite est distinct de l'estime de soi explicite. Finalement, des analyses exploratrices suggèrent qu'être conscient de l'accouplage n'était pas requis afin que le conditionnement survienne. Ces résultats de recherches appuient le modèle proposé de l'estime de soi implicite et fournissent une nouvelle tâche de conditionnement informatisée qui accroît l'estime de soi implicite.

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

Low self-esteem is associated with a host of costly phenomena, including depression, anxiety, aggression, drug and alcohol abuse, and general unhappiness (Diener & Diener, 1995; Leary, Schreindorfer, & Haput, 1995; see also Baumeister, 1993 for a review). In an attempt to understand some of the causes and consequences of low self-esteem, several decades of research have been devoted to uncovering the cognitive processes that underlie it. From this research, it is known that people with low self-esteem: have greater cognitive accessibility to personal weaknesses following failure than people with high self-esteem (Dodgson & Wood, 1998); overgeneralize the negative implications of failure in a specific domain to other, unrelated domains (Brown & Dutton, 1995); and show a distinct cognitive association between failure-rejection and success-acceptance, which is not evident in people with high self-esteem (Baldwin & Sinclair, 1996).

However, even with the depth of information available, a critical element of self-esteem may remain largely unexplored. In 2003 an issue of Psychological Inquiry was devoted to conceptualizing optimal self-esteem (Kernis, 2003). In this issue, self-esteem researchers from around the world discussed what it means to have high versus low self-esteem, the functions of self-esteem, the stability and insecurity of self-esteem, and the effects of an unconscious component of self-esteem called *implicit self-esteem*. It is this last topic, the unconscious component of self-esteem, which may be the critical unexplored element of self-esteem and which will be the focus of this dissertation.

Greenwald and Banaji (1995) first defined implicit self-esteem as the "introspectively unidentified (or inaccurately identified) effects of the self-attitude on evaluation of self-associated and self-dissociated objects" (p. 11). Essentially, implicit self-esteem is the automatic and unconscious component of self-esteem: a self-evaluation that occurs unintentionally and often outside of awareness (Farnham, Greenwald, & Banaji, 1999). Furthermore, there is evidence that implicit self-esteem taps into expressions of self-evaluation and feelings of self-worth that are distinct from traditional expressions of self-esteem (commonly referred to as explicit self-esteem) (Bosson, Swann, & Pennebaker, 2000). However, as implicit self-esteem is a relatively new concept, little is known about its cognitive development and representation.

The primary goal of the research presented in this dissertation was to uncover the cognitive mechanisms that contribute to the development and expression of implicit self-esteem. Drawing on past research supporting the interpersonal nature of self-esteem, I hypothesized that implicit self-esteem is rooted in an automatic and unconscious belief that the self is accepted or rejected by others, represented cognitively as associations between self and acceptance (self--*accepted)* and self and rejection (self--*rejected*). According to this conceptualization, high versus low implicit self-esteem is determined by the strength of the self--*accepted* cognitive association relative to the self--*rejected* association. Drawing from theories of learning, I further hypothesized that these associations are learned through evaluative conditioning. With repeated exposure, these associations become automatic and operate outside of awareness. Thoughts of the self, or exposure to objects related to the self, activate these self--*rejected* and self--*accepted* cognitive associations, which comprise a "gut-level" attitude toward the self.

To test this cognitive model, I developed a novel computer-based conditioning task that repeatedly paired self-relevant information with images of social acceptance. I hypothesized that participants who were exposed to this conditioning task would have higher levels of implicit self-esteem than participants who were exposed to a control version of the conditioning task.

The cognitive model of implicit self-esteem and the conditioning task used to enhance implicit self-esteem presented in this dissertation represent an original contribution to the self-esteem literature in general, and specifically to the understanding of implicit self-esteem. [1]

The Interpersonal Nature of Self-Esteem

I will argue that implicit self-esteem is represented as cognitive self--*accepted* and self--*rejected* associations that are formed through interpersonal experiences. This argument draws from past theory and research demonstrating that self-esteem, as it has been conceptualized in the past, is largely based on the social expectation that one is accepted or included. Specifically, low self-esteem stems, in part, from the explicit perception or expectation that one is rejected by others. By way of providing some background information, this section reviews some of the key theories, both historical and recent, which support an interpersonal model of self-esteem.

[1] At the same time that the research presented in this dissertation was being conducted, researchers at other Universities were exploring similar ideas (Dijksterhuis, 2004; Riketta & Dauenheimer, 2003). The results from their studies were published while data from Studies 2 and 3 of this dissertation were under review at Psychological Science (Baccus, Baldwin, & Packer, 2004). There are several key differences between Rikketa & Dauenheimer's (2003) and Dijksterhuis's (2004) research and the body of research presented in this dissertation. These differences will be addressed in Chapter 7, the General Discussion.

Two of the earliest scholars to study the self were Cooley (1902) and Mead (1934). Cooley (1902) theorized that self-esteem was largely interpersonal or social in nature, and is most well-known for the phrase "looking glass self." This expression refers to the idea that people construct their self by looking to society, or reactions from others, as their mirror. These *reflected appraisals* from others serve as a basis for self-evaluation, and are then incorporated into the self. Mead (1934) further theorized about the interpersonal nature of self-esteem with the concept of the *generalized other,* which represents a cumulative evaluation based on judgements and reactions from significant others in a person's life.

Cooley (1902) and Mead's (1934) focus on the social and interpersonal aspects of self-evaluation has carried into more recent conceptualizations of self-esteem. The most well-known of these theories is *sociometer* theory (Leary, Tambor, Terdal, & Downs, 1995b). Sociometer theory is based, in part, on the evolutionary standpoint that the need to belong is a basic motive (Leary & Baumeister, 2000). In the environment of evolutionary adaptedness, a human's chances of genetic survival (e.g., surviving until adulthood, successful mating) were greatly enhanced if that human was part of a group, especially if being alone meant having to compete against others for resources. Thus, the desire to belong and form meaningful social bonds might be an evolutionarily established basic need, along with the need for food, water, and shelter. Consequently, just as humans have a system to monitor thirst and hunger (i.e., a system that lets us know when it is time to consume food or water), sociometer theory contends that humans must also have a system for monitoring inclusion and exclusion. In sum, sociometer theory states that the self-esteem system functions as a gauge that "monitors the degree to which the

individual is being included versus excluded by other people and that motivates the person to behave in ways that minimize the probability of rejection or exclusion" (Leary, Tambor, et al., 1995, p. 518).

In a series of studies, Leary, Tambor, et al. (1995) provided support for this theory. In one study, participants who believed that they had been excluded from a group based on someone else's preferences rated themselves significantly more negatively than people who believed the selection had been randomly made. In another study, participants who received feedback that indicated rejection felt more negatively about themselves compared to participants who received either positive feedback or no feedback. Interestingly, those participants who received negative feedback tended to explicitly dismiss the feedback, that is, they rated it as inaccurate. Even though they believed that the exclusion was not justified or deserved, they still showed lower self-esteem than participants who were accepted. This demonstrates that people do not necessarily need to view exclusion as accurate or warranted in order for it to have a negative effect on self-esteem.

Finally, Leary, Tambor, et al. (1995) found a moderate correlation between self-esteem and perceived inclusion, suggesting that the relationship between trait self-esteem and perceived exclusion is likely reciprocal. The authors concluded that repeated experiences with social exclusion can lead to low self-esteem, and low self-esteem can lead to an increased predisposition to perceive exclusion.

Overall, this body of research provides strong evidence for the relationship between perceived social exclusion and self-esteem. Furthermore, over the past decade research based on sociometer theory has provided solid support for the interpersonal

nature of self-esteem (e.g., MacDonald, Saltzman, & Leary, 2003; Buckley, Winkel, & Leary, 2004).

While Leary, Tambor, et al.'s (1995) research has focused specifically on people's reactions to perceived exclusion, another line of research has focused on the cognitive mechanisms that underlie low versus high self-esteem. Baldwin (1992) theorized that people have cognitive *relational schemas*, which are cognitive representations of people's patterns of interpersonal interactions (Baldwin, 1992). A key component of the relational schema is the *interpersonal script*, which is conceptualized as an "if...then" pattern of relating based on repeated experiences of similar interactions (e.g., "If I cry, then my mom will hug me") (Baldwin, 1992). The relational schema, then, contains thoughts, feelings, and expectations about how another will respond to the self. Baldwin and Sinclair (1996) applied the relational schema to the study of self-esteem and showed that successes and failures are linked cognitively to the interpersonal outcomes of acceptance and rejection, respectively. Using a modified lexical decision task, Baldwin & Sinclair (1996) found that people with low self-esteem were faster to identify both positive social outcome words such as "cherished" when presented in the context of success, and negative social outcome words such as "despised" when presented in the context of failure, compared to people with high self-esteem. Thus, people with low self-esteem appear to possess cognitive associations between failure and rejection, suggesting that experiences involving failure lead them to expect rejection from others. People with high self-esteem did not show any evidence of a cognitive association between either failure-rejection or success-acceptance, suggesting that high self-esteem is related to expectations of interpersonal acceptance that are unrelated to personal successes or

failures. Furthermore, these findings were not found when the study was replicated using non-interpersonal words (e.g., freedom, amuse, stealing, decay).

In summary, research over the past several decades has provided support for an interpersonal theory of self-esteem, whereby people's explicit sense of self-worth is directly related to how accepted/included or rejected/excluded they feel. While there continues to be some debate surrounding the relationship between self-esteem and cognitive versus emotional reactions (e.g., Dutton & Brown, 1997), researchers increasingly agree that self-esteem is, in large part, an interpersonally based phenomenon related to cognitive representations of interpersonal acceptance and rejection.

Implicit Self-Esteem

In all of the aforementioned research, self-esteem was assessed and measured in explicit ways, most often using the Rosenberg Self-Esteem Scale (1965). The Rosenberg Self-Esteem Scale is comprised of ten statements including "I believe I am a person of worth" and "All in all, I am inclined to feel that I am a failure." When assessing state self-esteem, one of the most commonly used measures is the State Self-Esteem Scale (Heatherton & Polivy, 1991). The State Self-Esteem Scale contains 20 statements including "I feel that others respect and admire me" and "I feel inferior to others at this moment." Both the Rosenberg Self-Esteem Scale and the State Self-Esteem Scale are based on self-reports, which are subject to self-enhancing self-presentation strategies (Paulhus, 1986). In fact, research has indicated that high explicit self-esteem is positively correlated with a self-enhancing bias (Brown, 1991). Furthermore, Farnham et al. (1999) found that, for example, self-esteem measures inconsistently predict sensitivity to

feedback, predict behaviours that are more related theoretically to self-presentation strategies than to self-esteem, and predict self-enhancing behaviours mainly in situations where self-presentational demands are high. Thus, self-esteem as measured by self-report methods relies on the participants' explicit thoughts and feelings about the self that may be tainted by self-presentation strategies. However, measuring the unconscious component of self-esteem (i.e., implicit self-esteem) can help circumvent the effects of self-presentation that are evident in explicit self-esteem measurement.

Implicit self-esteem measurement instruments. Several researchers have developed various tools for assessing implicit self-esteem. In 2000, Bosson et al. examined the reliability and validity of seven implicit self-esteem measures, including the self-esteem Implicit Association Test (IAT) and the Name Letter Task. The self-esteem IAT is a computerized categorization task that measures automatic associations between *self-good* and *self-bad* by recording reaction times to target words coded as self-related (e.g., I, me), other-related (e.g., they, them), good (e.g., rainbow, happy), or bad (e.g., pain, vomit). For some of the trials, participants use the same computer key to categorize self-related words and good words. In a separate block of trials, participants use the same computer key to categorize self-related words and bad words. A difference score between these two blocks is calculated representing the strength of *self-good* versus *self-bad* associations.

In the Name Letter Task, participants are presented with all 26 letters of the alphabet and asked to rate, on a scale from 1 (*Dislike Very Much*) to 9 (*Like Very Much*) how much they like each letter. The task is presented as a "Linguistic Preferences Task" and participants are led to believe that their evaluation of the letters will be used to

develop stimuli for future studies. As such, participants are unaware that self-esteem is being assessed. Researchers have theorized that a preference for one's own name letter initials is an indication of high implicit self-esteem (Greenwald & Banaji, 1995; Jones, Pelham, Mirenberg, & Hetts, 2002; Kitayama & Karasawa, 1997; Koole & Pelham, 2003; Nuttin, 1985; 1987). Specifically, it has been theorized that individuals who possess positive self-feelings will transfer these positive evaluations onto objects that are associated with the self such as the letters of one's own name (Greenwald & Banaji, 1995).

In their review of implicit self-esteem measures, Bosson et al., (2000) found that the IAT and the Name Letter Task demonstrated satisfactory test-retest reliabilities. However, they also found that these two implicit measures of self-esteem did not show strong convergent validity, that is, they did not appear to tap into the same underlying construct. While this seems puzzling at first, researchers in the area of implicit cognition have recently highlighted that multiple measures of the same construct do not necessarily have to correlate, representing the multifaceted nature of implicit constructs (Cunningham, Preacher, & Banaji, 2001). This underscores the lack of knowledge surrounding the cognitive representation of implicit self-esteem. Indeed, while considerable effort has been devoted to developing ways to measure implicit self-esteem, little has been done to explore and uncover the cognitive makeup of implicit self-esteem.

What are the advantages of measuring implicit self-esteem? Assessing implicit self-esteem provides several benefits when compared to assessing explicit self-esteem alone. First, implicit self-esteem is measured indirectly, such that the participant is either unaware that any form of self-evaluation is being assessed, or unable to control his or her

responses (e.g., during reaction time tasks). Thus, the concerns regarding self-presentation effects outlined above are diminished when measuring implicit self-esteem, as unobtrusive measures are less subject to demand characteristics. Second, implicit self-esteem may provide previously unknown information regarding the cognitive mechanisms underlying self-esteem. The research of Baldwin and Sinclair (1996) presented earlier in the chapter provides evidence that differences in explicit self-esteem (i.e., low versus high) are mirrored in different cognitive representations. While Baldwin and Sinclair (1996) measured unconscious cognitive associations involving interpersonal events, rather than those that involved the self per se, their findings highlight the importance of considering automatic cognitive operations when assessing self-esteem. Third, and as I will discuss shortly, implicit self-esteem is not necessarily correlated with explicit self-esteem, suggesting that implicit and explicit self-esteem may represent distinct aspects of self-esteem. Implicit measures of self-esteem allow access to self-attitudes of which people may not be aware—that is, those that exist outside of conscious awareness. Research involving implicit self-esteem may provide information about self-evaluation that is not accessible through explicit measurement. The correlation or lack thereof between implicit and explicit self-esteem will be considered in more detail in the following section.

Finally, traditional efforts to increase self-esteem have tended to focus on altering the self-critical thoughts associated with low self-esteem (Brewin, 1989). That is, an explicit approach has been used. While deliberately addressing self-critical thoughts may influence explicit expressions of self-esteem, it is not known what influence this has, if any, on implicit self-esteem. Enhancing self-esteem at an explicit level alone may still

leave low self-esteem at an unconscious level. It seems necessary to address the underlying automatic negative self-thoughts (i.e., low implicit self-esteem) in order to effectively create positive self-esteem.

Implicit versus Explicit Self-Esteem

Implicit self-esteem is typically only weakly correlated with explicit self-esteem, if it is correlated at all (Bosson et al., 2000; Farnham et al., 1999; Greenwald & Farnham, 2000). That is, a person might self-report positive self-evaluations, yet at an unconscious level have negative self-esteem. This indicates that implicit and explicit self-esteem are psychologically distinct. Researchers have explored the relationship between implicit and explicit self-esteem, and their findings are reviewed below. Following this review, I will introduce a model of dual attitudes that can be applied to self-esteem to help explain the lack of correlation between implicit and explicit self-esteem, as well as to support the presumption that implicit and explicit self-esteem operate using different cognitive systems.

Explicit self-esteem can be measured as a trait (e.g., using the Rosenberg Self-Esteem Scale (1965)) or as a state (e.g., using the Heatherton and Polivy State Self-Esteem Scale (1991)). Some research has shown that implicit self-esteem can also be influenced on a momentary basis. For example, participants who were given the opportunity to engage in self-affirmation of a subjectively important value following failure showed more positive scores on the Name Letter Task compared to participants who did not engage in self-affirmation (Koole, Smeets, van Knippenberg, & Dijksterhuis, 1999). Furthermore, studies have shown that reminders of personal death lead to a

decrease in name letter evaluations (Koole, Decehesne, & van Knippenberg, 2000, cited in Koole, Dijksterhuis, & van Knippenberg, 2001). Finally, while not in the domain of self-esteem, research has shown that the IAT can be used to capture changes in automatic attitudes. Dasgupta and Greenwald (2001) used a racial IAT to assess automatic racial attitudes and found that exposure to admired Black and disliked White exemplars led to a weakened pro-White attitude. This provides evidence not only for the malleability of automatic attitudes, but also for the use of the IAT to capture any such changes. Overall, these findings demonstrate that the lack of correlation between implicit and explicit self-esteem is not simply due to shifting levels of explicit self-esteem across a stable level of implicit self-esteem.

Several studies have examined the predictive power of implicit versus explicit self-esteem. Explicit self-esteem was found to be a significant predictor of subjective well-being, while implicit self-esteem was uncorrelated with subjective well-being (Schimmack, & Diener, 2003). However, in other studies implicit self-esteem was found to be better than explicit self-esteem at predicting, for example, observer-rated nonverbal anxiety and negative mood in response to threatening feedback, as well as persistence at tasks and socially undesirable behaviour in response to failure (Bosson et al., 2000; Greenwald & Farnham, 2000; Jordan, Spencer, & Zanna, 2002; Spalding & Hardin, 1999).

Researchers have also examined the specific implications of possessing congruent (i.e., low implicit and low explicit self-esteem or high implicit and high explicit self-esteem) versus incongruent self-esteem (i.e., high explicit and low implicit self-esteem or low explicit and high implicit self-esteem). Particular attention has been paid to the

combination of high explicit/low implicit self-esteem, referred to as *defensive high self-esteem* by some researchers (Brown & Bosson, 2001; Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003). Research has shown that people with defensive self-esteem show higher levels of narcissism, demonstrate greater levels of in-group bias (a form of self-enhancement), and engage in more dissonance reduction in relation to personal choice compared to people high in both implicit and explicit self-esteem (Jordan et al., 2003). Other research has suggested that people high in explicit and low in implicit self-esteem show more unrealistic optimism, more strongly prefer positive to negative personality descriptions, and report smaller self-ideal discrepancies than people who possess congruent implicit and explicit self-esteem (Bosson, Brown, Zeigler-Hill, & Swann, 2003).

Furthermore, Shimizu and Pelham (2004) found that, consistent with past research, both implicit and explicit self-esteem were independently related to positive life events. In line with past research showing that positive life events can be psychologically disruptive for some people, Shimizu and Pelham (2004) found that for people low in either implicit or explicit self-esteem, positive life events were associated with illness. For people high in implicit or explicit self-esteem, positive life events were not significantly related to illness. When the congruence of implicit and explicit self-esteem was tested, it was found that there was no relationship between positive life events and illness for people with a combination of high implicit/low explicit self-esteem. However, positive life events were associated with illness for people with the combination of high explicit/low implicit self-esteem. This finding suggests that, in this instance, the influence of implicit self-esteem overrides the influence of explicit self-esteem. While the results of

these studies shed some light on the interaction between implicit and explicit self-esteem, the implications of possessing congruent versus incongruent self-esteem are still relatively unclear.

To address theoretically the lack of correlation between implicit and explicit self-esteem, I turn to a model from the attitudes literature. Recently, a model of dual attitudes has been proposed suggesting that it is possible for people to simultaneously hold two different attitudes toward a single attitude object—one at a conscious, deliberative, explicit level, and the other at an automatic, unconscious, implicit level (Wilson, Lindsey, & Schooler, 2000). This model of dual attitudes has been applied to self-esteem to propose that implicit and explicit self-esteem co-exist cognitively as two distinct self-evaluations that are not necessarily correlated (Jordan et al., 2003; Koole et al., 2001). Wilson et al.'s (2000) dual attitudes model challenges old models which posit that when people change or develop a new attitude, it replaces the old attitude in memory. In a dual attitudes model, it is argued that the first attitude is not replaced, but remains in memory as part of a dual attitude. Applied to self-esteem, both implicit and explicit self-evaluations coexist in memory: the implicit attitude is activated automatically, whereas the explicit attitude requires conscious thought. Furthermore, even when the explicit attitude is activated, the implicit attitude can continue to have an unconscious influence on implicit and uncontrollable responses.

Overall, implicit and explicit self-esteem appear to be distinct and independent of each other. A dual attitudes approach to self-esteem sheds some light on the representation of self-esteem at an implicit versus explicit level, suggesting that it is possible to possess conflicting self-evaluations at an automatic versus a deliberative level.

Summary of implicit self-esteem research to date. When looking at the current state of implicit self-esteem research as a whole, one can see that there has been a considerable amount of research devoted to developing measures of implicit self-esteem and exploring the relationship between implicit and explicit self-esteem. From this research, only a few broad and general conclusions about implicit self-esteem can be drawn: 1) implicit self-esteem is an automatic and unconscious evaluation of the self; 2) implicit and explicit self-esteem are not necessarily correlated; and 3) implicit and explicit self-esteem do not necessarily relate to behaviour in the same fashion. This leaves a large gap in our understanding of implicit self-esteem. I have introduced a dual attitudes model to help explain the independence of implicit and explicit self-esteem, and to introduce the possibility that implicit and explicit self-esteem operate using different cognitive systems. However, this model does not fully explain how implicit self-esteem is represented cognitively, how it is learned, or how it can be changed. I begin to address these issues in the next section by turning to theories of learning.

Theories of Learning

Classical conditioning is most commonly explained using Pavlov's historical experiment. By pairing a tone with food, Pavlov soon had dogs salivating in response to the presentation of the tone alone (Pavlov, 1927). Early theories of classical conditioning focused on the physiological and mechanical processes, as opposed to the informational or cognitive processes, involved in the creation of an association between two events (Pavlov, 1927; Rescorla & Wagner, 1972; see also Domjan, 2000 for a review). Specifically, classical conditioning was originally viewed as stimulus-response (S-R) learning. An unconditioned stimulus (US; e.g., food) elicits an unconditioned response

(UR; e.g., salivation). A conditioned stimulus (CS; e.g., tone) is paired with the UC so that eventually, the presentation of the CS alone leads to a conditioned response (CR; e.g., salivation in response to the tone even without presentation of the food). More recent interpretations of classical conditioning focus on the learning of an association between the CS and the US, referred to as stimulus-stimulus (S-S) learning (Rescorla, 1988). Furthermore, a type of classical conditioning that involves the associative transfer of valence, called *evaluative conditioning*, has been explored as a means to influence the development of likes and dislikes. I theorized that the principles of evaluative conditioning could be applied to implicit self-esteem in order to develop and test a cognitive model of implicit self-esteem. Before expanding on the details of this application, I will review some of the basic principles of evaluative conditioning.

Evaluative conditioning specifically refers to "changes in the liking of a stimulus that result from pairing that stimulus with other positive or negative stimuli" (De Houwer, Thomas, & Baeyens, 2001, p. 853). A common evaluative conditioning paradigm involves pairing pictures rated as "neutral" by participants with other pictures that the participant had previously rated as "liked" or "disliked." Following this pairing, participants tend to rate the neutral picture according to the affective valence of the picture with which it was paired.

Evaluative conditioning does not appear to rely on contingency awareness. Unlike Pavlovian conditioning, which generally occurs only when participants are aware of the contingency between the CS and the US (for a review, see Dawson & Schell, 1987), a review by De Houwer, Baeyens, and Hendrickx (1997; cited in De Houwer et al., 2001) suggests that evaluative conditioning is independent of contingency awareness. It is

important to note that there are two types of awareness: demand awareness and contingency awareness. Demand awareness refers to an awareness of the experimental hypothesis as reported by the participant, while contingency awareness refers to an awareness of the relationship between the CS and the US as reported by the participant (Field, 2000). Several studies have shown that evaluative conditioning can be obtained even in the absence of demand awareness or contingency awareness (e.g., Baeyens, Eelen, & Van den Bergh, 1990; Hammerl, Bloch, & Silverthorne, 1997).

Evaluative conditioning appears to be highly resistant to extinction (for a review see De Houwer et al., 2001). For example, in one study, neutral pictures were paired with either positive or negative pictures. Following the expected evaluative conditioning effect, participants were exposed to unreinforced presentations of the neutral stimuli. Participants continued to evaluate the neutral pictures based on what these pictures had been paired with in the learning stage, demonstrating a lack of extinction (Baeyens, Eelen, Van den Bergh, & Crombez, 1989).

Evaluative conditioning also demonstrates sensitivity to counterconditioning procedures (De Houwer et al., 2001). A counterconditioning procedure is used to eliminate a previously conditioned response by pairing the CS that elicited the response with a new US that elicits a different response (Lieberman, 1993). Baeyens et al. (1989) demonstrated that counterconditioning could be effective in evaluative conditioning paradigms. In their study, participants were conditioned to like or dislike previously neutral pictures (determined through a baseline rating) through an evaluative conditioning procedure. Neutral pictures were paired with pictures that the participant had previously rated as either liked or disliked. After this phase, some of the participants were exposed

to a counterconditioning procedure, whereby the same previously neutral pictures (now rated positively or negatively) were paired with pictures of a valence opposite to that previously paired (e.g., pictures that had been paired with liked pictures during the initial conditioning phase were now paired with disliked pictures). Results showed that although participants developed an affective preference for the neutral photographs through the initial evaluative conditioning procedure, this preference was removed if they were subsequently exposed to counterconditioning. Furthermore, there was evidence that the counterconditioning procedure caused the neutral picture to acquire the affective valence of the *opposite* sign (e.g., from liked to disliked, or from disliked to liked).

These findings suggest that human evaluative learning is sensitive to current information in the environment, particularly if the new information is in the form of a contingency opposite to one previously learned (Baeyens et al., 1989). It is important to note that this change in affective preference was found even when participants reported no contingency awareness. These findings also indicate that counterconditioning might be a useful procedure to employ when trying to change maladaptive or problematic responses to certain stimuli.

Finally, De Houwer et al. (2001) argue that conceptually, evaluative conditioning could be considered a type of referential learning, rather than Pavlovian signal or expectancy learning. In Pavlovian conditioning, the presentation of the CS (e.g., tone) leads to an explicit expectancy that the US (e.g., food) is actually going to occur in the near future. The strength of the link between the CS and the US depends on the extent to which the CS is a reliable and nonredundant predictor of the US. If the US does not occur, extinction of the CS-CR link is likely to occur. Conversely, in evaluative

conditioning the presence of the CS (e.g., picture) leads to the conscious or unconscious activation of the US (e.g., a liked picture), without involving any explicit expectancy that the US is actually going to occur. Even if the US does not occur, the CS will continue to elicit the CR.

While this referential model is currently proposed as more functional than computational, De Houwer et al. (2001) liken the process of referential learning to that of a simple Hebbian learning rule. The strength of association between two stimuli is increased when those two stimuli are presented together; however, the association strength remains intact even when one of the two stimuli is presented in isolation. The advantage of a referential system such as this is that when a stimulus is encountered, its valence is automatically determined by "averaging across the valence of the stimuli with which the stimulus co-occurred in the past" (De Houwer et al., 2001, p. 865).

Applying Evaluative Conditioning to Implicit Self-Esteem

Drawing from theories of learning, and specifically De Houwer et al.'s (2001) referential model of evaluative conditioning, I theorized that implicit self-esteem is acquired through evaluative conditioning. Repeated pairing of the self and social rejection or acceptance transfers the valence of the affect associated with "being accepted" or "being rejected" to the self. This eventually creates a self--*accepted* or self--*rejected* cognitive link. Presentation of a self-relevant object (e.g. one's first name), or thoughts about the self automatically and unconsciously activate this cognitive association, leading to the experience of low or high implicit self-esteem. Finally, I also theorized that implicit self-esteem should be susceptible to counterconditioning, that is,

pairing self with acceptance should lead to the creation of a self--*accepted* link, even if the existing cognitive association is one of self--*rejected.*

Thus, by applying the principles of evaluative conditioning to implicit self-esteem I developed an interpersonal cognitive model of implicit self-esteem. I next developed a methodology for testing this hypothesized model. Considering that the model was based on the principles of evaluative conditioning, I decided to create an evaluative conditioning paradigm. While applying the principles of evaluative conditioning to an interpersonal cognitive model of implicit self-esteem is novel, the application of learning theories within social psychology is not. Before outlining the conditioning task used to test the hypothesized cognitive model of implicit self-esteem, I review a summary of research where the principles of classical conditioning have been successfully applied to social psychological phenomena.

Applying Theories of Learning to Social Psychology

Social psychologists have been applying the principals of conditioning to their research for years, beginning with the simple pairing of attitude objects with positive or negative words. This research discovered that pairing nonsense words with either positively or negatively toned words led participants to transfer an affective preference for the nonsense word based on whether it had been paired with a positive or negative word (Staats & Staats, 1958). More modern applications of conditioning have often centred on conditioned preference and aversion, that is, manipulating affective responses to an object (Johnsrude, Owen, Zhao, & White, 1999). Social psychologists have also focused on the role of conditioning in attitude formation (for a review see: Olson &

Fazio, 2002; Walther, 2002; Zanna, Kiesler, & Pilkonis, 1970). For example, the persistence of prejudiced attitudes is thought to be a result of unconscious pairings of "non-white" with negative stereotypes, which continue to be portrayed by the media. This unconscious evaluation can exist even in the face of a conscious acceptance of non-whites (e.g., Dovidio & Gaertner, 1998).

Olson and Fazio (2001) demonstrated a novel and strong effect of attitude conditioning. A key feature of their paradigm was the absence of contingency awareness, that is, participants were not aware of the pairing between the attitude object and the negative or positive stimulus. This is of particular importance because when there is contingency awareness in a conditioning study, there is an increased risk of demand characteristics. In Olson and Fazio's (2001) study, the novel attitude objects were various Pokemon™ cartoon creatures. Participants were shown a stream of unrelated words and images on a computer screen. The words and images were sometimes presented alone, and sometimes in pairs. Throughout the presentation, and in the experimental condition, a specific Pokemon™ character was paired with a positive item, while a different character was paired with a negative item. In Study 1, participants who saw the character paired with positive items evaluated that same character post-manipulation more positively than they did the character paired with negative items. No participants were able to explicitly recall which characters were paired with which items. In Study 2, Olson and Fazio (2001) used an implicit measure of attitudes (the IAT) following the conditioning manipulation. The same effect of conditioning was found, with participants who saw a positive item with the Pokemon™ character showing an implicit preference for that same character post-manipulation, compared to their automatic preference for the character paired with

negative items. These findings demonstrate the effectiveness of conditioning in the area of attitude formation.

Finally, conditioning has been used to help socially anxious individuals deal with a stressful social interaction. In a novel design, Baldwin and Main (2001) created associations between feelings of social acceptance or rejection and previously neutral environmental cues. To do this, they created a phoney questionnaire during which female participants were told that they would be given feedback about whether or not their responses were "socially desirable." In reality, the feedback was presented randomly for all participants in order to create associations between computer tones and social acceptance or rejection. Sometimes a row of smiling photographs was presented while one distinct computer tone was played. Other times a row of frowning photographs was presented while a second distinct computer tone was played. Following the questionnaire, participants were told that they would be having a short "get to know you" conversation with another student, who was actually a confederate of the experiment. This male confederate was instructed to act in a cool and aloof manner in order to create an anxiety provoking interaction. During the interaction, a computer on the other side of the room repeatedly emitted one of the two tones that had been conditioned during the questionnaire portion of the study. Results showed that, in general, participants who heard the computer tone that had been paired with smiling faces reported less anxiety during the interaction compared to participants who heard the computer tone that had been paired with frowning faces.

In summary, traditional theories of learning have been applied to social psychological phenomena such as the formation of attitudes and social anxiety. The

success of conditioning paradigms in other areas of social psychology indicated that a conditioning paradigm would be an effective way to test my hypothesized cognitive model of implicit self-esteem.

Conditioning Implicit Self-Esteem: Overview of Studies and Hypothesis

To review, drawing from theories of learning, and specifically De Houwer et al.'s (2001) referential model of evaluative conditioning, I hypothesized that implicit self-esteem is developed through evaluative conditioning. Repeated exposure to social rejection or acceptance creates a transfer of the affect associated with "being accepted" or "being rejected" to the self. This eventually creates a self--*accepted* or self--*rejected* cognitive link. The interpersonal aspect of implicit self-esteem (i.e., rooted in acceptance/rejection from others) draws from theory and research showing that explicit self-esteem is strongly related to interpersonal inclusion and exclusion (e.g., Cooley, 1902; Baldwin & Sinclair, 1996; Leary, Tambor, et al., 1995). These self--*rejected* or self--*accepted* cognitive associations are activated through presentations of the self or self-relevant information.

To test this hypothesized interpersonal cognitive model of implicit self-esteem, and based on past success with conditioning methodologies in social psychology, I developed a computer-based conditioning task that repeatedly paired the self (in the form of self-relevant information) with social acceptance (in the form of photographs of smiling faces). While the notion of pairing self-relevant information (e.g., first name) with photographs of smiling faces to enhance implicit self-esteem may seem overly simple, if, as hypothesized, implicit self-esteem is based in cognitive self--*accepted* and

self--*rejected* associations learned through evaluative conditioning, then directly exposing participants to an evaluative conditioning procedure that repeatedly pairs the self with acceptance should enhance implicit self-esteem.

Overview of studies. Four studies are presented in this dissertation. In Study 1, I developed a computer-based conditioning task to test the basic hypothesis that pairing self-relevant information with photographs of smiling faces would enhance implicit self-esteem (as measured by the IAT). In Study 2, participants' mood was measured following the conditioning task to confirm that the conditioning task was not simply leading to enhanced mood. In Study 3, a pre-measure of implicit self-esteem was included to determine if pre-manipulation levels of implicit self-esteem influenced the effect of the conditioning task. Furthermore, I attempted to assess contingency awareness by including a post-manipulation question asking participants to guess how often their self-relevant information was paired with photographs of smiling faces. Finally, in Study 4 the number of men included in the sample was increased to gain generalizability, and the question used to assess contingency awareness was refined. Finally, supplementary analyses are presented in Chapter 6. In this chapter I present post-hoc analyses that explore the effects of the conditioning task in more detail, and on issues that are less central to the thesis. Discussed are the results of a meta-analysis, findings from pooled data sets, and results from a second measure of implicit self-esteem (the Name Letter Task).

CHAPTER 2

CONDITIONING IMPLICIT SELF-ESTEEM: STUDY 1

The main purpose of Study 1 was to design and test a computer-based conditioning task for enhancing implicit self-esteem. The goal of the conditioning task was to create an association between self and interpersonal acceptance. To do this, I developed a computer-based conditioning task in which participants' self-relevant information (e.g., first name, birthday) was repeatedly paired with social acceptance, presented in the form of photographs of smiling faces. I hypothesized that participants who completed this version of the conditioning task would show higher levels of implicit self-esteem compared to those who completed a control version of the task.

Method

Participants

Participants were forty students (18 men, 18 women, 4 unspecified) from McGill University in Montreal. Age was not recorded; however, all participants were undergraduate students. Participants received course credit for their participation.

Materials

Measures of explicit self-esteem. Explicit self-esteem was measured pre-manipulation using the Rosenberg Self-Esteem Scale (Rosenberg, 1965). The Rosenberg Self-Esteem Scale contains ten items (e.g., "I feel I am a person of worth", "At times I think I am no good at all") and is scored on a 7-point Likert scale (1 = *Strongly Disagree*, 7 = *Strongly Agree*).

Explicit self-esteem was measured post-manipulation using the State Self-Esteem Scale (Heatherton & Polivy, 1991). The State Self-Esteem Scale contains twenty items and participants are asked to "circle the response that best represents your thoughts RIGHT NOW." Choices are: *not at all*, *a little bit*, *somewhat*, *very much*, and *extremely*.[2]

Conditioning Task. The conditioning task was programmed using E-Prime software (2002). Instructions appeared onscreen for participants to follow. Participants were first asked to enter in six self-relevant pieces of information using the computer keyboard. These items were: first name, birthday, hometown, ethnicity, phone number, and street name. These items were a sample of those used in an idiographic self-esteem IAT designed by Greenwald & Farnham (2000). Following this, participants were instructed that they would see a word appear in one of four quadrants on the computer screen. They were instructed to use their computer mouse to click on this word as quickly as possible, and informed that clicking on the word would cause an image to briefly appear in the quadrant where the word had appeared. Participants were informed that the image would disappear and a different word would appear in another quadrant; they were instructed to again click on this word. They were instructed to go as quickly as possible and to try not to make any mistakes. However, they were also told that making a few mistakes was natural, as they were trying to go quickly. If participants made a mistake and clicked on an empty box, the phrase "Oops, you missed! Please continue." was shown on the screen.

---

[2] Other post-questionnaires were included in the studies within this dissertation, however no significant effects were found with them so they will not be discussed. The one post-measure that was included in all studies was the Self-Liking/Competency Scale (SLCS; Tafarodi & Swann, 1995). Exploratory regression analyses were conducted including the SLCS score as a predictor of implicit self-esteem and no significant effects were found, that is, any differences in implicit self-esteem between conditions were not due to changes in post-measured self-liking or self-competency.

The words that appeared in the boxes were either self-relevant (i.e., entered in by the participant at the beginning of the task) or non-self-relevant. Non-self-relevant words were programmed into the task as generic lists of first names, dates, cities, ethnicities, phone numbers and street names. In the experimental version of the task, participants' self-relevant information was always followed by an image of a smiling face. A frowning or a neutral face followed non-self-relevant items. In the control condition, both self-relevant and non-self-relevant information were followed by a random selection of smiling, frowning, and neutral faces (see Appendix 1 for a depiction of the conditioning task). All face photographs appeared on the screen for 400 ms. In total, participants were presented with 240 trials, 80 of which were self-relevant. Participants in both conditions saw the three expressions presented an equal number of times, and saw the same number of self-relevant versus non-self-relevant information. Crucially, participants in the experimental and control conditions saw an equal number of smiling faces (see Table 1 for a breakdown of the word-face pairings in both conditions).[3]

The digital colour face photographs were drawn randomly from a database in the laboratory, which contained equal numbers of male and female faces. While the majority of faces were Caucasian, a variety of ethnicities were included in an attempt to represent the ethnic make-up of the study body at McGill University. The photographs were taken

[3] In the literature on classical conditioning, the question of how to construct a control condition has been debated (for a discussion see Domjan, 2000, pp. 54-57; also Cacioppo, Marshall-Goodell, Tassinary, & Petty (1992)). In particular, Domjan noted, "a universally applicable and acceptable solution to the control problem in ... conditioning is not available." (2000, p. 54). For the present studies, it seemed most critical to control for the number of smiles presented in the control versus the experimental condition. Thus, the control condition consisted of non-contingent pairings of the US and CS (i.e., self-relevant information was paired with accepting faces some of the time, frowning faces some of the time, and neutral faces some of the time. Furthermore, non-self-relevant information was sometimes paired with accepting faces). This type of control group has been used in a previous study involving the classical conditioning of attitudes (Cacioppo et al., 1992).

from the undergraduate population at McGill University and thus were mostly of young individuals between the ages of 18 and 23.

EXPERIMENTAL CONDITION				
	Smile	Frown	Neutral	Total Words
Self-Relevant Words	80	0	0	80
Non-Self-Relevant Words	0	80	80	160
Total Faces	80	80	80	Total Trials: 240

CONTROL CONDITION				
	Smile	Frown	Neutral	Total Words
Self-Relevant Words	26-27	26-27	26-27	80
Non-Self-Relevant Words	53-54	53-54	53-54	160
Total Faces	80	80	80	Total Trials: 240

Table 1: A breakdown of the pairings between target words and face photographs in the experimental and control conditions of the computer-based conditioning task.

The Self-Esteem Implicit Associations Task (IAT). As discussed earlier, the self-esteem IAT was developed by Greenwald & Farnham (2000) and is based on the original IAT by Greenwald, McGhee, & Schwartz (1998). For the present study, the target words were taken from a list developed and used by Greenwald & Farnham (2000) in their study demonstrating the validity of the self-esteem IAT (see Appendix 2). What follows is a detailed explanation of the IAT.

The standard procedure of the IAT was followed for this and all subsequent studies. There are five blocks of trials in the IAT, two of which are critical blocks of trials used to calculate the participant's IAT score. In the first two blocks, participants

categorize target words into one of two categories. In the first block participants categorize self- or other-related target words into the categories of *"Self"* or *"Other"* (e.g., me, them, I, their). In the second block participants categorize pleasant- and unpleasant-meaning target words into the categories of *"Good"* or *"Bad"* (e.g., rainbow, vomit, sunshine, pain). Next is the first of two critical blocks of trials, where participants categorize all four types of target words as either *"Self or Good"* using one computer key, or as *"Other or Bad"* using a different computer key (see Figure 1). In this block, for example, participants use the same computer key to categorize "me" and "rainbow."

In the final critical block of trials, participants categorize target words as *"Other or Good"* using one computer key, or as *"Self or Bad"* using a different computer key (i.e., the reverse of the first critical block; see Figure 2). In this block, for example, participants use the same computer key to categorize "me" and "vomit."

Both critical blocks consisted of 10 practice trials and 20 test trials. To calculate a score on the IAT, reaction times between the two critical blocks are compared. If a person shows a large difference between the two critical blocks, it suggests that they were faster to categorize the words in one block compared to the other. The faster the participant categorizes the target words, the stronger the association between the paired concepts is presumed to be. Thus, participants who have slower reaction times when *"Self"* and *"Bad"* are categorized using the same computer key compared to when *"Self"* and *"Good"* share a computer key are presumed to have high levels of implicit self-esteem. These participants find it easier to think of *"Self"* and *"Good"* together than to think of *"Self"* and *"Bad"* together.

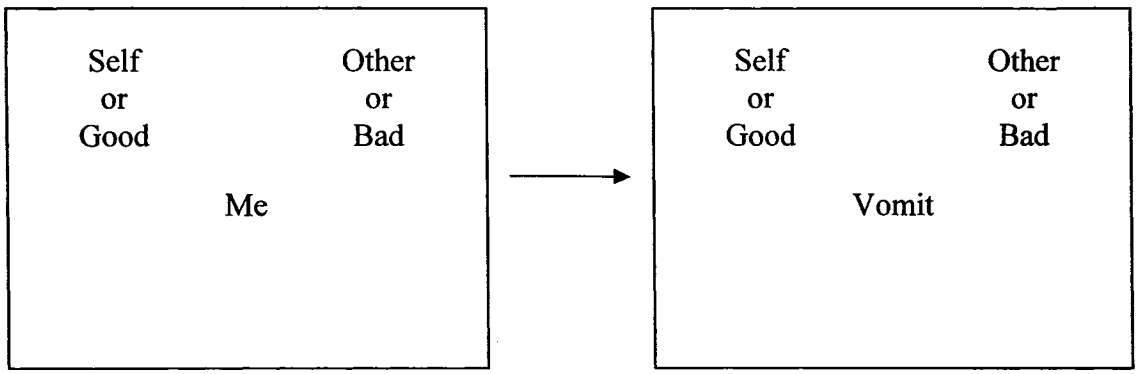


Figure 1. First critical block of the self-esteem IAT. Participants are presented with target words in the middle of the screen. They are asked to categorize them as "Self or Good" using the "e" key on the computer keyboard or as "Other or Bad" using the "i" key on the computer keyboard.

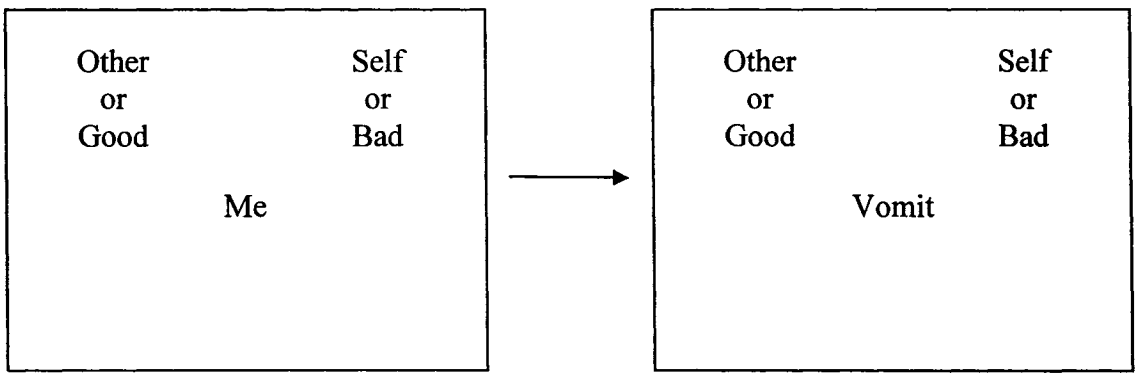


Figure 2. Second critical block of the self-esteem IAT. Participants are presented with target words in the middle of the screen. Compared to the first critical block, the positions of "Self" and "Other" have been reversed. Participants are now asked to categorize the target words as "Other or Good" using the "e" key on the computer keyboard or as "Self or Bad" using the "i" key on the computer keyboard.

To encourage participants to strive for accuracy, a green 'O' was displayed on the screen for 300 ms after each correct response and a red 'X' was displayed on the screen for 300 ms after each incorrect response. After an incorrect response, the IAT required participants to enter the correct response in order to continue with the next target word.[4]

Procedure

Participants believed they were taking part in a reaction time study called "How Fast Are You?" which measured personality variables and reaction times. They were not aware that the study was related to self-esteem. As participants arrived for the study they were seated at a computer. Up to six participants were tested at one time. All participants first completed a standard consent form and were assigned a subject identification number.

The testing session began with participants completing the Rosenberg Self-Esteem Scale. Following this, they were randomly assigned to either the experimental or control version of the conditioning task and asked to follow the instructions on the screen. All participants then completed the IAT. After completion of both computer tasks, participants were given the post-questionnaire package that included the State Self-Esteem Scale. The entire session lasted between 30-45 minutes.

---

[4] Some additional data were collected in Studies 1,2, and 4 using an IAT that included a different order for the presentation of the two critical blocks, which is not common in research using the self-esteem IAT. In this order, the first critical block consisted of *"Self"* and *"Bad"* sharing a computer key, while the second critical block consisted of *"Self"* and *"Good"* sharing a computer key. In initial validation tests of the IAT, Greenwald et al. (1998) found some small effects of order such that larger IAT effects were observed when the first critical block presented the combination that tends to be easier for most subjects (e.g., self-good). In tests of the self-esteem IAT, Greenwald and Farnham (2000) found this similar order effect, with larger effects being evident when the *"Self or Good"* block was presented first. Furthermore, many researchers currently do not adopt a counterbalancing method of critical block order when using the self-esteem IAT (Bosson et al., 2000; Jordan et al., 2003; Karpinski & Hilton, 2001). Consistent with the data presented in this dissertation, researchers tend to present the *"Self or Good"* critical block first, followed by the *"Self or Bad"* critical block. In the current data, results were inconclusive for participants who completed the order where *"Self"* and *"Bad"* were paired together in the first block. For this reason, and because the norm in the literature is to present the *"Self or Good"* critical block first, these additional data were excluded from the analyses.

At the end of the study participants were fully debriefed as to the nature of the study, thanked, and dismissed.

Analytic Strategy

Regression Analyses

Multiple regression was used to analyze the data from this and all subsequent studies. While the main hypothesis was that condition would predict post-measured implicit self-esteem scores, I also wanted to control for the possible influence of pre-measured self-esteem. As such, following Aiken and West (1991), I centred the pre-measured explicit self-esteem score from the Rosenberg Self-Esteem Scale. Also following Aiken and West (1991), I then calculated a cross-product interaction term between condition and explicit self-esteem (in Studies 3 and 4 the pre-measured implicit self-esteem score from the Name Letter Task was also centred and cross-product interactions were calculated and included in the regression analyses). In the event that the higher-order interactions were not significant, I again referred to Aiken and West (1991) to interpret lower order effects. They note that if the higher order interaction is not significant, the researcher should select an analytic strategy according to whether an interaction is expected according to theory. Specifically, they write, "If...there is not a strong theoretical expectation of an interaction, a step-down procedure should be used, where the non-significant higher order interaction is dropped from the equation." (p. 105). Because my main hypothesis did not include the expectation of an interaction, I adopted an exploratory approach for subsequent analyses by dropping the non-significant higher order interactions from the equation.

Calculation of the IAT Score

Based on an algorithm developed by Greenwald et al. (1998), response latencies longer than 3,000 ms were recoded as 3,000 ms and response latencies shorter than 300 ms were recoded as 300 ms. The average reaction times to self-related targets from the first critical block (i.e., when "*Self*" and "*Good*" shared a computer key) were subtracted from the average reaction time to self-related targets from the second critical block (i.e., when "*Self*" and "*Bad*" shared a computer key). Higher values represent higher implicit self-esteem.

In the traditional calculation of the self-esteem IAT, responses to both self-related and other-related targets are included in the calculation of the score. According to this calculation, a high score suggests that the participant has strong associations between both self-related words and good words, and other-related words and bad words. However, researchers have highlighted the importance of considering the specific role of the "other" in the IAT and have used alternate ways of calculating the IAT score to test this (Karpinski, 2004; Karpinski & Hilton, 2001). Furthermore, Greenwald, Banaji, Rudman, Farnham, Nosek, and Mellott (2002) note that there are alternative data management procedures for the calculation of the IAT score that do not produce substantively different results from those obtained with their suggested algorithm.

In the present research, I was most concerned with the effect of the conditioning task on self-related targets. I did not expect the conditioning task to have an influence on other-related words—that is, I did not expect the conditioning task to influence participants' reaction times to other-related target words in the *"Other or Good"* versus *"Other or Bad"* blocks of trials. For these reasons, primary analyses involving the IAT

were conducted using a score based on self-targets only. Results using other calculations of the IAT are reported in the preliminary analyses.

Results

Preliminary Analyses

Because I used a non-traditional calculation of the IAT score, I conducted several preliminary regression analyses using other calculations of the IAT score. All regressions considered condition, pre-measured explicit self-esteem, and the interaction between the two as predictors. First, I carried out a regression analysis on the traditional calculation of the IAT score, where both self-related and other-related targets are included in the score. No significant effects of condition were found.[5] I then ran a regression analysis on an IAT score calculated using only the other-related target words and again, no significant effects were found. This is particularly important considering that the experimental condition also includes the pairing of non-self-relevant information (e.g., "other") with photographs of neutral and frowning faces. The null finding from reaction times of IAT other-related targets alone confirms that the conditioning task was not creating an enhanced "other-bad" association.

Finally, I conducted a regression analysis on a baseline reaction time IAT score. This variable was calculated by computing the mean reaction time of all target words other than self-related (i.e., "other", "good", and "bad" words) in the two critical blocks. No significant main effect of condition was found, suggesting that participants did not tend to respond faster or slower overall depending on which condition they were in. In all subsequent regression analyses involving the IAT, the IAT baseline reaction time

[5] In this case, and in the case of all other non-significant results reported in this dissertation, $p > .10$.

variable was included to control for individual differences in the general tendency to respond faster or slower on the task.

Correlations between implicit and explicit self-esteem

As expected, pre-measured (Rosenberg Self-Esteem Scale) and post-measured (State Self-Esteem Scale) explicit self-esteem were correlated, $r = .60, p < .001$. Unexpectedly, and inconsistent with past research (Bosson et al., 2000), pre-measured explicit self-esteem was correlated with post-measured implicit self-esteem (as measured by the IAT), $r = .32, p = .04$. Furthermore, post-measured explicit self-esteem was also correlated with post-measured implicit self-esteem, $r = .43, p = .005$.

Implicit Self-Esteem

I regressed condition, explicit self-esteem, and the cross-product interaction between the two on the IAT score (self-targets only, as in all analyses to be discussed). There was a significant interaction between condition and pre-measured explicit self-esteem, $\beta = -.60$, $t(38) = -2.34, p = .025$. Following Aiken and West (1991), to explore this interaction, I tested simple slopes at values one standard deviation above and below the mean of implicit and explicit self-esteem. Results showed that participants who began the study low in explicit self-esteem benefited the most from the conditioning task, $\beta = .57$, $t(38) = 2.75, p = .009$ (see Figure 3). Thus, the computer-based conditioning task enhanced implicit self-esteem for participants who began the study with low explicit self-esteem.

There were no significant effects involving gender.

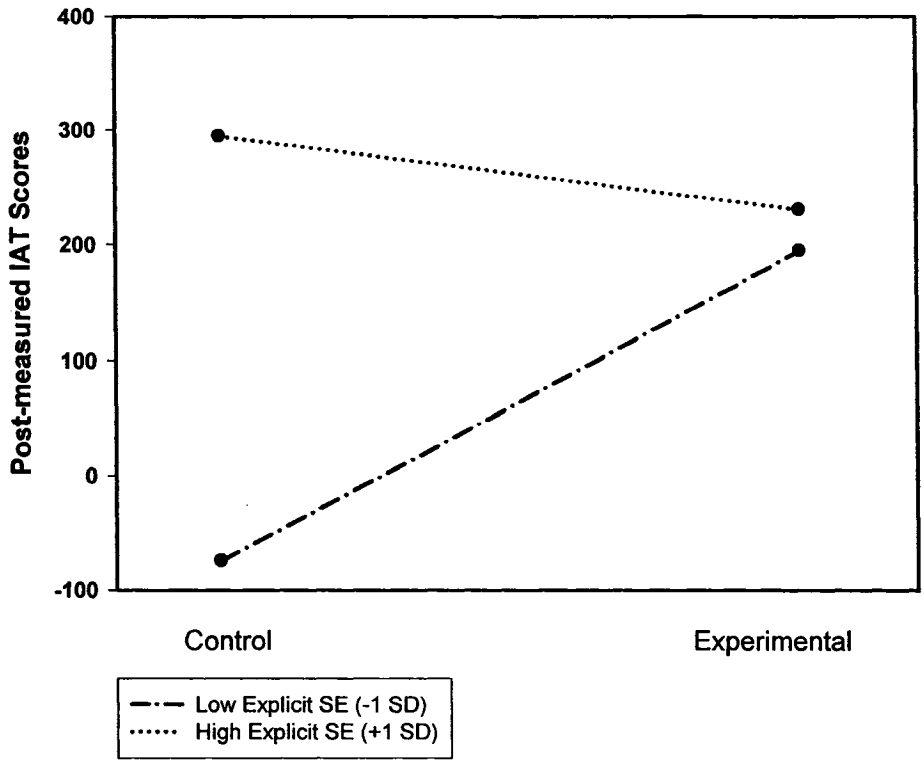


Figure 3. Study 1: Post-measured implicit self-esteem as a function of condition and pre-measured explicit self-esteem. SE = Self-esteem; SD = Standard Deviation.

Explicit Self-Esteem

I conducted a regression analysis on post-measured explicit state-self-esteem with pre-measured explicit self-esteem and the interaction between pre-measured explicit self-esteem and condition included in the equation. There were no significant effects involving condition. Thus, pairing self-relevant information with photographs of smiling faces had no influence on self-esteem as measured through explicit self-report, even though explicit self-esteem was correlated with implicit self-esteem.

Discussion

The results of Study 1 provided some support for the hypothesis that implicit self-esteem is represented cognitively as a link between "self" and "accepted" or "self" and "rejected." Specifically, the two-way interaction between condition and pre-measured explicit self-esteem showed that for participants who began the study low in explicit self-esteem, the computer-based conditioning task led to enhanced implicit self. While this latter result is unexpected, it seems reasonable that participants who began the study low in self-esteem benefited the most from the conditioning task. However, there is no clear theoretical explanation for this interaction.

Finally, Study 1 also showed that the conditioning task had no effect on post-measured explicit self-esteem. The lack of findings on explicit self-esteem suggest two things: 1) demand characteristics were not a factor, as participants did not appear to be consciously reporting enhanced self-evaluations in the experimental group; and 2) the fact that participants showed change at an implicit level but not at an explicit level is in line with a dual attitudes model of self-esteem. A dual attitudes model argues that even when explicit responses are retrieved from memory, implicit attitudes can still influence implicit responses (Wilson et al., 2001). Thus, even though participants may have been self-reporting negative self-evaluations, positive implicit self-esteem was being demonstrated through their automatic responses (i.e., their reaction times on the IAT).

In summary, Study 1 provided some evidence that pairing self-relevant information with social acceptance in the form of photographs of smiling faces has a positive influence on implicit self-esteem. However, both the interaction and the correlation between implicit and explicit self-esteem were unexpected. Furthermore, the

possibility remains that pairing self-relevant information with photographs of smiling faces leads to enhanced mood, which in turn accounts for the tendency to associate "self" and "good" on the IAT. In Study 2, I included a post-manipulation measure of mood to control for the effects of mood on the relationship between the conditioning task and enhanced implicit self-esteem.

CHAPTER 3

CONDITIONING IMPLICIT SELF-ESTEEM: STUDY 2

The main goal of Study 2 was to determine whether or not the conditioning task influenced participants' mood. If enhanced implicit self-esteem is simply a by-product of positive mood, then participants who complete the experimental version of the conditioning task should show higher levels of positive affect compared to participants in the control condition. To test this, I included a post-manipulation measure of mood. I hypothesized that creating a self--*accepted* cognitive association does not lead to enhanced mood. I further hypothesized that participants in the experimental condition would show higher IAT scores compared to participants in the control condition; however, the two conditions would not differ in terms of positive affect.

Method

Participants

Participants were twenty-one students (3 men, 18 women) from McGill University in Montreal. Age was not recorded; however, all participants were undergraduate students. Students received partial course credit for their participation.

Materials and Procedure

The materials and procedure were the same as Study 1 with the following exception:

Profile of Mood States (POMS). The depressed/elated, anxious/calm, and confident/unconfident subscales of the POMS (Lorr, & McNair, 1982) were included as a post-measure to determine whether or not enhanced implicit self-esteem is related to

enhanced mood. The POMS is a list of 36 adjectives (e.g., "cheerful", "timid", "calm") that participants are to rate on a scale from zero (*much unlike this*) to 3 (*much like this*) according to how much the word describes how they are feeling at that moment.

Results

Preliminary Analyses

As in Study 1, I ran several preliminary regression analyses using other calculations of the IAT score. All regressions considered condition, pre-measured explicit self-esteem, and the interaction between the two as predictors. First, I ran a regression analysis on the traditional calculation of the IAT score, where both self-related and other-related targets are included in the score. No significant effects were found. I also ran a regression analysis on an IAT score calculated using just the other-related target words and again, no significant effects were found. Finally, I ran a regression analysis on the baseline reaction time IAT score, with no significant effects. The IAT baseline reaction time variable was included in subsequent regression analyses.

Correlations between implicit and explicit self-esteem

Consistent with past research (Bosson et al., 2000), pre-measured explicit self-esteem was not correlated with post-measured implicit self-esteem, $r = .04$, *ns*. Furthermore, post-measured explicit self-esteem was not correlated with the IAT, $r = .25$, *ns*. Pre-measured and post-measured explicit self-esteem were marginally correlated, $r = .38$, $p = .054$. Thus, the correlation between implicit and explicit self-esteem found in Study 1 was not evident here.

Mood

I hypothesized that the conditioning task would not have an effect on mood. Preliminary regression analyses confirmed that condition (either alone or interacting with pre-measured explicit self-esteem) was not a significant predictor of post-manipulation mood. Thus, participants who saw their self-relevant information paired with photographs of smiling faces did not report significantly higher scores on the POMS mood scale compared to participants in the control condition. The centred mood score from the POMS was included in all subsequent analyses to control for any effects of increased mood on the overall model.

Implicit Self-Esteem

I regressed condition, explicit self-esteem, and the cross-product interaction between the two on the IAT score. The two-way interaction from Study 1 was not significant in this sample, thus the interaction term was dropped from the equation.

Condition had a marginally significant main effect on IAT scores, $\beta = .29$, $t(16) = 2.07$, $p = .055$, indicating that people in the experimental condition had higher scores on the IAT compared to those in the control group. Recall that the IAT score represents the discrepancy between reaction times when *"Self"* and *"Good"* were paired together compared to when *"Self"* and *"Bad"* were paired together. Participants in the experimental group had a mean of 307.83 on the IAT while those in the control group had a mean of 195.29, showing that participants in the experimental condition had faster reaction times (i.e., were quicker to categorize self-related words) when *"Self"* and *"Good"* were paired together compared to when *"Self"* and *"Bad"* were paired together.

Thus, pairing self-relevant information with images of social acceptance enhanced implicit self-esteem.

There was an insufficient number of men to conduct any analyses on gender.

Explicit Self-Esteem

A regression analysis was conducted on post-measured explicit state-self-esteem with pre-measured explicit self-esteem and the interaction between pre-measured explicit self-esteem and condition included in the equation. There were no significant effects. As in Study 1, pairing self-relevant information with photographs of smiling faces did not influence explicit self-reports of self-esteem.[6]

Discussion

Study 2 found evidence that the conditioning task led to higher implicit self-esteem for participants in the experimental condition compared to those in the control condition. Participants who were repeatedly exposed to the pairings of their self-relevant information with photographs of accepting faces appeared to have an easier time associating "self" with "good" than did those participants who saw their self-relevant information paired with a random selection of smiling, frowning, and neutral faces.

The main goal of Study 2 was to confirm that the conditioning task was specifically influencing cognitive self--*accepted* associations as opposed to simply enhancing mood. Indeed, results showed that mood did not differ between participants who had completed the experimental versus control version of the conditioning task. This shows that higher scores on the IAT following the conditioning task are unrelated to higher scores on the measure of mood.

[6] The lack of statistical effects on mood and explicit self-esteem must be interpreted carefully, as non-significant results do not necessarily imply an absolute lack of effects. However, the consistency across studies suggest that the conditioning task was not influencing self-reports of mood or explicit self-esteem.

The lack of results involving post-measures of explicit self-esteem again suggests that while the conditioning task has a positive effect on implicit self-esteem, it has no influence on explicit self-esteem. While participants in the experimental condition demonstrated enhanced implicit self-evaluations, they did not self-report more positive self-evaluations than those in the control condition.

Studies 1 and 2 provided evidence that the computer-based conditioning task led to enhanced implicit self-esteem. However, the interaction between condition and pre-measured explicit self-esteem found in Study 1 was not evident in Study 2. Thus, it remained unclear whether or not the interaction in Study 1 was spurious, or if the effect of the conditioning task is actually influenced by pre-measured self-esteem.

CHAPTER 4

CONDITIONING IMPLICIT SELF-ESTEEM: STUDY 3

Studies 1 and 2 demonstrated that evaluative conditioning influences implicit self-esteem, and that this effect is not influenced by mood. This finding provides support for the hypothesized interpersonal cognitive model of implicit self-esteem. The main goal of Studies 3 and 4 was to explore some of the finer details of the effect of the conditioning task on implicit self-esteem.

In Study 3, I sought to assess the influence of pre-measured self-esteem on the effects of the conditioning task. An interaction effect was found in Study 1; however, a main effect was found in Study 2. Furthermore, the significant correlation between implicit and explicit self-esteem in Study 1 implies that participants who began the study with low explicit self-esteem may also have had low implicit self-esteem. It is possible that the conditioning task may only be effective for participants who have low implicit self-esteem to begin with, suggesting that for those who begin the study with high implicit self-esteem, exposure to self--*accepted* pairings does not significantly influence their implicit self-evaluations compared to participants who are not exposed to such pairings. However, because there was no pre-measure of implicit self-esteem in Study 1 or in Study 2, it was impossible to statistically measure this theoretical possibility. Thus, in Study 3 I wanted to control for the possibility that the conditioning task might affect participants differently depending on their pre-manipulation level of implicit self-esteem. I decided not to give participants the IAT as a pre-measure of implicit self-esteem as I was concerned that repeated exposure to the IAT might lead to practice effects.

Consequently, I included the Name Letter Task as both a pre-manipulation and post-manipulation measure of implicit self-esteem.

A second goal of Study 3 was to test the effect of demand characteristics influencing the results. Although the lack of effects on explicit self-esteem in both Studies 1 and 2 suggests that demand characteristics were not a factor, there was no explicit measure of contingency awareness. In Study 3 I included a question at the end of the testing session asking participants to estimate how often their self-relevant information was paired with photographs of smiling faces during the computer-based conditioning task.

Method

Participants

Participants were forty-nine undergraduate students (9 men, 40 women) from McGill University in Montreal. Age ranged from 18-25 with a mean of 19.5. Participants received either partial course credit or $8.00 for their time.

Materials and Procedure

The materials and procedure were the same as Study 2 with the following exceptions:

Name Letter Task. The Name Letter Task was included both pre- and post-manipulation. As discussed earlier, the Name Letter Task asks participants to rate all 26 letters of the alphabet according to how much they like them on a scale from 1 (*dislike very much*) to 9 (*like very much*). The task was presented to participants under the guise of collecting materials for future studies on linguistic preferences. For the post-measure version of the

Name Letter Task, participants were instructed to "not look back at your previous ratings of these letters."

Assessing awareness of the contingency. In order to control for the possibility that demand characteristics might have been present if participants in the experimental condition were aware of the pairing between their self-relevant information and photographs of smiling faces, I included a question at the end of the post-manipulation questionnaire package to measure participants' awareness of the contingency. The question read as follows:

> Think back to the computer task you completed, where you clicked on words that appeared in boxes. A picture of a face followed these words. Some of the words you clicked on were, in fact, words that you provided before the task began (i.e., your first name, your birthday). When you clicked on the words you provided during this computer task, approximately how often do you recall them being following by a smiling face?
>
> Please circle a percentage below, OR, place an "X" in the box if you do not remember how often the words you entered were followed by a positive image:

0%	10 - 20%	30 - 40%	50%	60 - 70%	80 - 90%	100%
never	almost never	less than half of the time	half of the time	more than half of the time	almost always	always

OR

☐ I do not remember how often the words I entered were followed by a smiling face (please place an "X" in the box)

Results

Preliminary Analyses

Preliminary regression analyses again confirmed that condition (either alone or interacting with pre-measured implicit or explicit self-esteem) was not a significant predictor of scores on the post-manipulation measure of mood. Mood was included as a variable in subsequent regression analyses to control for the possible influence of mood on the overall model.

Preliminary regression analyses on other calculations of the IAT score mirrored those of Studies 1 and 2. No significant effects were found in regression analyses of the traditional calculation of the IAT score or on an IAT score calculated using just the other-related target words. Finally, a regression on the baseline reaction time IAT score did not show any significant effects. In all subsequent regression analyses involving the IAT, the IAT baseline reaction time variable was included in the equation.

Calculation of the Name Letter Score

To control for response styles involving the tendency to use high or low numbers on the scale, I first ipsatized ratings by subtracting from each participant's rating of his or her initials the mean liking score the participant gave to the remaining letters of the alphabet. To then control for a potential confounding tendency of individuals to rate certain frequently used letters of the alphabet higher than other, less frequent letters (Jones, Pelham, & Mirenberg, 2002), I subtracted from the rating of each of these letters the mean ipsatized score for all other participants who did not share that initial. Each participant's score was the mean of the adjusted rating for his or her two initials.[7]

[7] Results from analyses involving the post-measured Name Letter scores were generally in line with findings from the IAT and suggested that the experimental version of the conditioning task was

Correlations between implicit and explicit self-esteem

Pre-measured explicit self-esteem was not significantly correlated with the IAT, or the Name Letter scores ($rs < .10$, *ns*). Furthermore, post-measured explicit self-esteem was not correlated with the IAT or the Name Letter scores ($rs < -.22$, *ns*). Pre-measured and post-measured explicit self-esteem were significantly correlated ($r = .75, p < .001$).

Implicit Self-Esteem

Regression analyses yielded no significant interactions between condition, pre-measured explicit self-esteem or pre-measured implicit self-esteem, thus all interaction terms were dropped from the regression.

Condition had a significant main effect on IAT scores, $\beta = .32$, $t(47) = 2.33$, $p = .024$. Replicating the finding from Study 2, participants in the experimental condition had higher scores on the IAT compared to those in the control group. Those in the experimental group had a mean of 220.42 on the IAT while those in the control group had a mean of 133.57. Pre-measured implicit self-esteem had a marginal main effect on post-measured IAT scores, $\beta = .26$, $t(47) = 1.87$, $p = .068$, suggesting that participants who began the study high in implicit self-esteem had high implicit self-esteem post-manipulation. Pre-measured implicit self-esteem did not interact with condition.

Gender did not appear to have any effect; however, there were again an insufficient number of men to test for gender effects in the data.

---

enhancing participants' name letter ratings. However, regression analyses involving the Name Letter scores were less consistent than analyses involving the IAT. Discussion regarding the Name Letter findings is left to the supplementary analyses section of this dissertation (Chapter 6). Interested readers are also invited to refer to Baccus et al. (2004) for a discussion of post-manipulation findings on the Name Letter Task.

Exploratory Interaction Analyses

While the interaction between condition, pre-measured implicit self-esteem, and pre-measured explicit self-esteem was not significant, for the purpose of comparison with Study 4, I created a graphical representation of how participants who were incongruent in implicit/explicit self-esteem (i.e., low explicit/high implicit or high explicit/low implicit) reacted to the conditioning task compared to those who were congruent in implicit/explicit self-esteem (i.e., high explicit/high implicit or low explicit/low implicit). Following Aiken & West (1991), values at one standard deviation above and below the mean of implicit and explicit self-esteem were graphed. Figure 4 shows that all four combinations of implicit/explicit self-esteem appeared to benefit from the effects of the conditioning task compared to those in the control condition.

Explicit Self-Esteem

There were no significant effects on explicit self-esteem. As in Studies 1 and 2, the conditioning task had a positive effect on participants' automatic self-evaluations, yet it did not influence their explicit self-evaluative reports.

Assessing Awareness of the Contingency

Unexpectedly, the majority of the participants in the experimental condition (n = 14 out of a total of 25) selected the "I do not remember how often the words I entered in were followed by a smiling face." Of the remaining 11 participants in the experimental condition, only 3 of these participants correctly guessed that their self-relevant information was paired with smiling faces 100% of the time. This suggests that participants in the experimental condition were generally unaware of the 100% contingency between their self-relevant information and smiling faces; however, due to

the large proportion of "I do not remember" responses an exact conclusion could not be made.

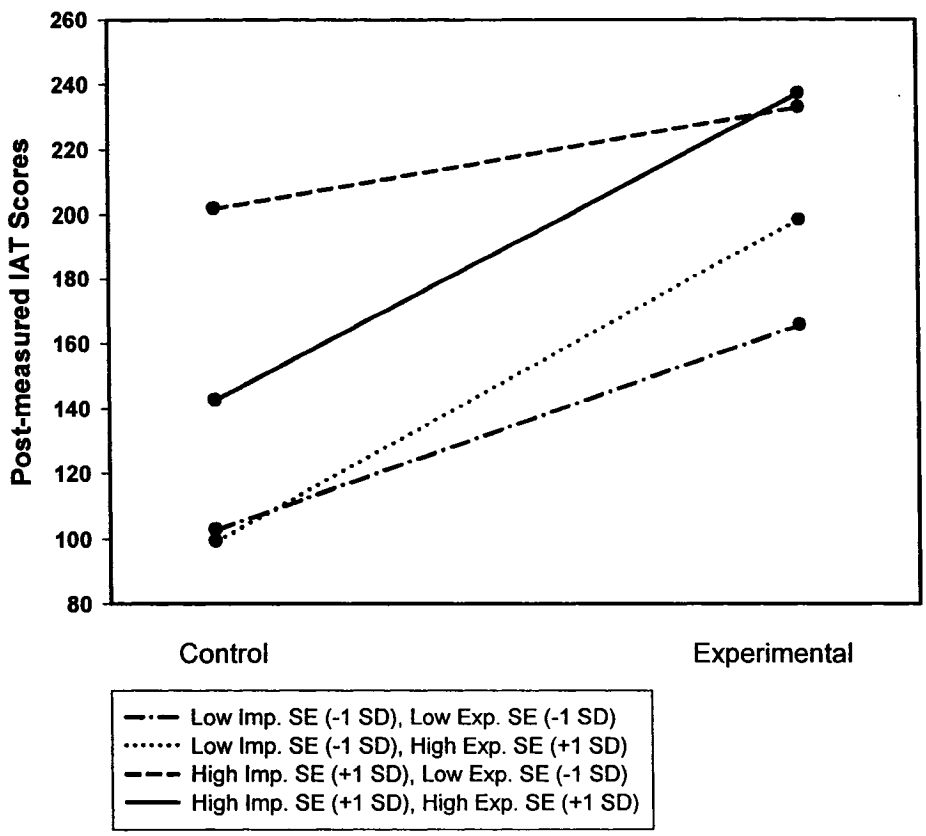


Figure 4. Study 3: Post-measured implicit self-esteem as a function of condition, pre-measured implicit self-esteem and pre-measured explicit self-esteem. SE = Self-esteem; SD = Standard Deviation; Imp = Implicit; Exp = Explicit.

Discussion

Study 3 replicated the main finding from Study 2. The computer-based conditioning task had a positive effect on implicit self-esteem. Participants who completed the experimental version of the task showed higher scores on the IAT compared to participants in the control condition. The lack of interaction effects in

Studies 2 and 3 suggest that the 2-way interaction in Study 1 may have been a spurious effect. Finally, the conditioning task did not influence post-measured explicit self-esteem, replicating the finding from Studies 1 and 2.

In general, participants in the experimental condition did not appear to be aware of the contingency between their self-relevant information and photographs of smiling faces. Unfortunately, a thorough analysis of participants' awareness of the contingency was not possible due to the majority of participants selecting the "I do not remember" option when asked to guess how often their self-relevant information was paired with photographs of smiling faces. Due to this, a firm conclusion could not be determined. Specifically, it is possible that some participants were aware of the contingency, yet were not confident of their guess and thus selected the "I don't remember" option. That is, some participants may have decided to select the "I don't remember" option rather than select the "wrong" percentage. The main goal of Study 4 was to determine whether or not demand characteristics were influencing the results.

CHAPTER 5

CONDITIONING IMPLICIT SELF-ESTEEM: STUDY 4

The results from Study 3 replicated the main effect from Study 2, namely that pairing self-relevant information with photographs of smiling faces led to enhanced implicit self-esteem. Pre-measured self-esteem did not have any influence on this finding. However, some unresolved issues remained. One concern was the imbalance between men and women in the sample. Due to the constraints of the subject pool used to recruit participants, both Studies 2 and 3 did not have a large sample of men, thus it was difficult to determine if there were any effects of gender. No significant gender effects were found in Study 1; however, no pre-measure of implicit self-esteem was included in that study. I increased the number of men in the sample for Study 4. I also sought to further explore the issue of contingency awareness. In Study 3, it was unclear whether or not contingency awareness occurred, due to the fact that most participants selected the "I do not remember" option when guessing how often their self-relevant information was paired with smiling faces. In Study 4, I removed that option and forced participants to make a percentage selection when guessing the contingency between their self-relevant information and photographs of smiling faces.

Method

Participants

Participants were sixty-nine students (27 men, 42 women) from McGill University and Dawson College in Montreal. Age ranged from 16-29 with a mean of 18.8. Participants received either partial course credit or $8.00 for their time.

Materials and Procedure

The materials and procedure were the same as Study 3 with one exception: *Assessing awareness of the contingency.* In Study 4, the "I do not remember" option was removed and participants were forced to guess a percentage.

Results

Preliminary Analyses

Due to a procedural error, 11 participants did not complete the POMS (mood measure). Preliminary regression analyses on those participants who did complete the POMS again confirmed that condition (either alone or interacting with pre-measured self-esteem) was not a significant predictor of changes in mood. Due to the procedural error, and considering the lack of effects of mood in Studies 2 and 3, mood was not included in subsequent analyses.

As in the previous studies, I ran regression analyses using alternate calculations of the IAT score. No significant effects were found on the IAT score calculated using other-targets alone, or on the traditional calculation of the IAT (i.e., including both self and other target words in the calculation). Finally, a regression on the baseline reaction time IAT score did not show any significant effects. In all subsequent regression analyses involving the IAT, the IAT baseline reaction time variable was included.

Correlations between implicit and explicit self-esteem

As in Studies 2 and 3, pre-measured explicit self-esteem was not significantly correlated with the IAT, or the Name Letter scores ($rs < .18$, *ns*). Furthermore, post-measured explicit self-esteem was not correlated with the IAT or the Name Letter scores

($rs < .16$, *ns*). Pre-measured and post-measured explicit self-esteem were significantly correlated ($r = .67, p < .001$).

Implicit Self-Esteem

I conducted a regression analysis of the IAT score including condition, pre-measured explicit self-esteem, pre-measured implicit self-esteem, and all possible interactions as predictors. The IAT baseline variable was included in the regression equation.

I found only an unexpected significant three-way interaction between condition, pre-measured implicit self-esteem and pre-measured explicit self-esteem, $\beta = .47$, $t(60) = 2.37$, $p = .02$. Condition alone was not a significant predictor, nor were the 2-way interactions. According to this 3-way interaction, post-measured IAT scores depended on participants' level of implicit and explicit self-esteem prior to the beginning of the task, and on condition. To explore this interaction, I tested simple slopes at values one standard deviation above and below the mean of implicit and explicit self-esteem (Aiken & West, 1991). As shown in Figure 5, participants who began the study high in both explicit and implicit self-esteem (solid line) and then completed the experimental version of the conditioning task showed significantly higher post-measured IAT scores compared to their counterparts in the control condition, $\beta = .60$, $t(60) = 2.25$, $p = .028$. Unexpectedly, participants who began the study high in explicit self-esteem but low in implicit self-esteem had lower IAT scores in the experimental condition (dotted line), compared to their counterparts in the control condition, $\beta = -.51$, $t(60) = -2.12$, $p = .038$. The simple slopes analysis for participants who began the study with congruent low self-esteem approached significance ($p = .11$, Mean IAT Control = 140.87. Mean IAT Experimental

= 228.24), and the simple slopes analysis for participants who began the study with high implicit/low explicit self-esteem was not significant, $p = .49$.

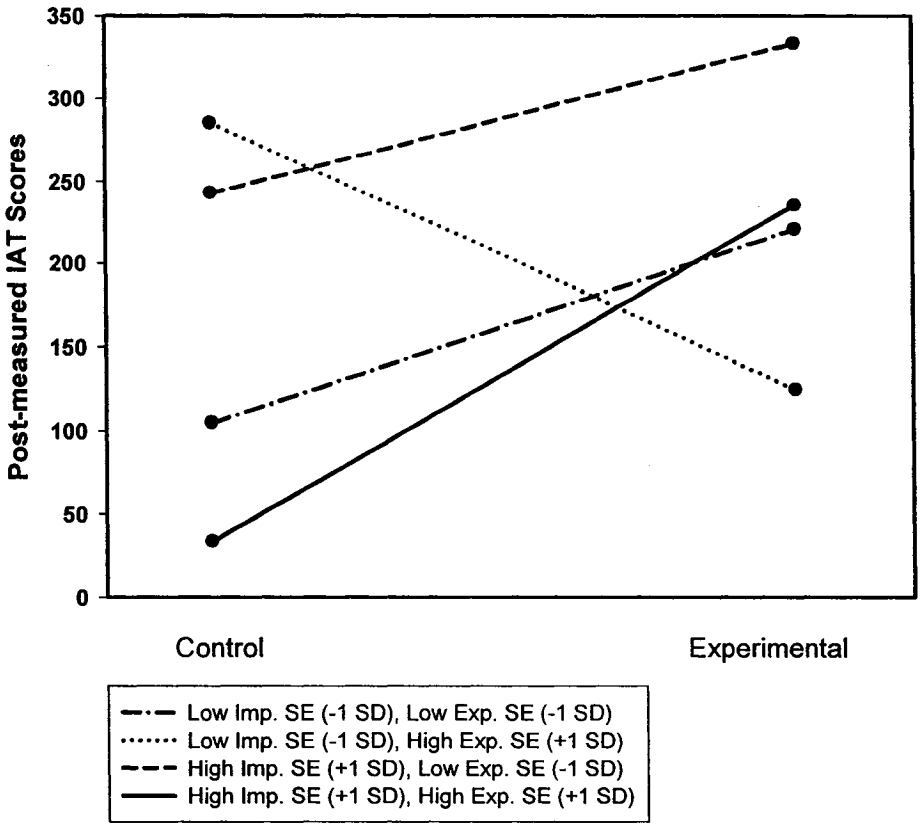


Figure 5. Study 4: Post-measured implicit self-esteem as a function of condition, pre-measured implicit self-esteem and pre-measured explicit self-esteem. SE = Self-esteem; SD = Standard Deviation; Imp = Implicit; Exp = Explicit.

Exploratory Analyses on Gender

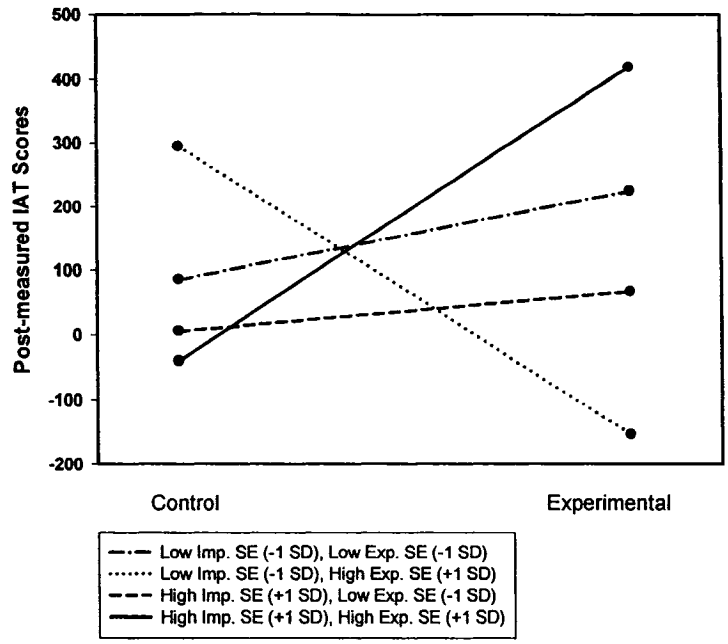
A regression analysis did not show any significant interactions involving gender. However, considering the low sample size and the corresponding lack of power to detect a 4-way interaction, I conducted some exploratory analyses. Splitting the data file on gender and running the regression model revealed a significant main effect of condition

in women, $\beta = .37$, $t(37) = 2.58$, $p = .014$. Women in the experimental condition had a mean IAT score of 230.62, while those in the control condition had a mean of 151.56. Analysis of men alone revealed a marginally significant three-way interaction, $\beta = .93$, $t(18) = 1.93$, $p = .07$. An exploratory simple slopes analysis revealed that, in line with the finding from the full data set above, men who began the study high in both implicit and explicit self-esteem had higher IAT scores compared to their counterparts in the control condition, $\beta = .75$, $t(18) = 1.78$, $p = .09$. Furthermore, those participants with defensive self-esteem (high explicit/low implicit) had lower IAT scores in the experimental condition compared to those in the control condition, $\beta = -95$, $t(18) = -2.63$, $p = .017$ (see Figure 6a). Thus, the three-way interaction appeared to be driven by the reactions of men who began the study with defensive self-esteem. Although not significant, the three-way interaction in women was graphed for comparison (see Figure 6b).

Explicit Self-Esteem

A regression analysis was conducted on post-measured explicit self-esteem with condition, pre-measured explicit self-esteem, pre-measured implicit self-esteem, and all possible interaction terms included in the equation. As in Studies 1, 2, and 3, no significant effects of condition were found, demonstrating that pairing self-relevant information with photographs of smiling faces does not influence explicit reports of self-esteem.

(a)



(b)

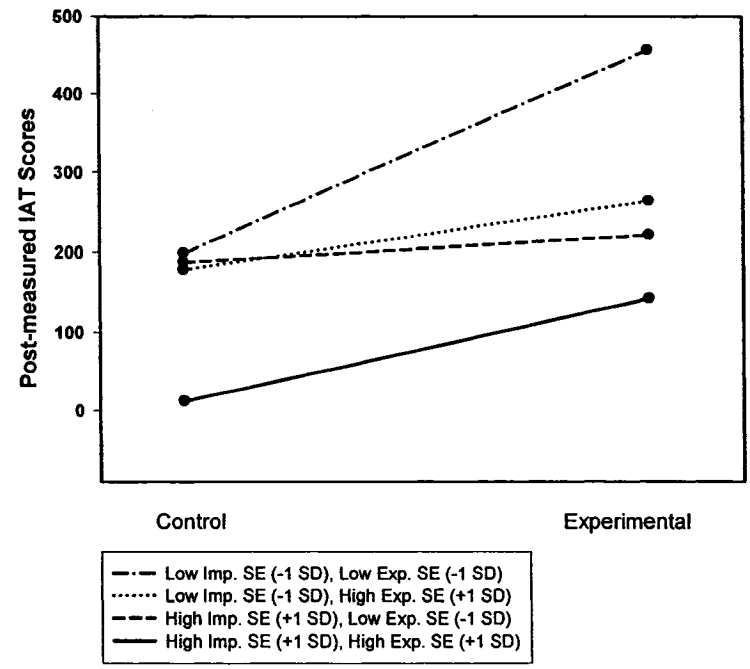


Figure 6. Study 4: Post-measured implicit self-esteem as a function of condition, pre-measured implicit self-esteem and pre-measured explicit self-esteem in men (a) and women (b). SE = Self-esteem; SD = Standard Deviation; Imp = Implicit; Exp = Explicit.

Contingency Awareness

Answers to the question asking participants to guess the percentage of time that their self-relevant information was paired with smiling faces was coded as follows: 0% = 1, 10-20% = 2...100% = 7. Interestingly, no participants in the experimental condition guessed that there was a 100% contingency between presentation of their self-relevant information and photographs of smiling faces (i.e., a 7 on the contingency awareness measure) confirming that in general, participants in the experimental condition were not aware of the 100% contingency. Results of an ANOVA showed that there was no significant difference between the experimental group and the control group on their estimates of any contingency between their self-relevant information and photographs of smiling faces, $F(1, 65) = .163$, *ns*. Participants in the experimental condition guessed a mean of 3.74 on the scale, compared to a mean of 3.92 in the control condition, indicating that participants in both groups guessed that their self-relevant information was paired with photographs of smiling faces on average approximately 50% of the time. This strongly suggests that overall, participants in the experimental condition were not aware of the 100% pairing between their self-relevant information and photographs of smiling faces.

My initial reason for including a question assessing participants' awareness of the contingency was to ensure that demand characteristics were not influencing the conditioning effect. This was shown to be the case. However, past research has explored the role of contingency awareness in conditioning to see if contingency awareness is necessary for learning to occur (De Houwer, 2001). I ran exploratory analyses to determine the exact role of contingency awareness in this paradigm, specifically, to

address whether or not contingency awareness matters when applying evaluative conditioning to implicit self-esteem.

For the first analysis, I excluded participants who indicated that they were aware of the contingency in the experimental condition. Because I was interested specifically in identifying those who were primarily aware of the 100% contingency (a median split would include participants who guessed a 60% contingency), I split the data file at the 75th percentile on the frequency distribution of the contingency awareness variable (M = 5.5). This excluded participants who had selected a 6 on the contingency awareness variable, which corresponded to guessing that their self-relevant information was paired with photographs of smiling faces greater than 80% of the time. Eight participants were excluded from the analysis. Results showed the same three-way interaction found with the full data set, confirming that demand characteristics did not play a role in the outcome.

To explore the effect of contingency awareness further, I specifically compared those participants in the experimental group who were aware of the contingency to those who were not. To do this, I created one group consisting of participants in the experimental condition who had selected a 6 on the contingency awareness measure (i.e., most aware of the contingency; $n = 6$) and the other group consisted of participants in the experimental condition who had selected a 1 on the contingency awareness measure (i.e., least aware of the contingency; $n = 6$). Using t-tests, I compared the means between these two groups to each other, as well as to the control group (n = 37). Results showed that the only two groups that differed were the contingency unaware group when compared to the control group, $t(41) = 1.89$, $p = .088$. While this effect is marginal, and the sample size is

low, this finding implies that a lack of contingency awareness may actually be a contributing factor to the effectiveness of the conditioning task on implicit self-esteem. The role of contingency awareness will be addressed further in the General Discussion.

Discussion

As in the previous studies, the computer-based conditioning task influenced implicit self-esteem. However, an unexpected three-way interaction was found between condition, pre-measured implicit self-esteem and pre-measured explicit self-esteem. An analysis of this interaction found that participants who began the study with high congruent self-esteem (i.e., high in both implicit and explicit self-esteem) benefited the most from the conditioning task. Surprisingly, participants who began the study with defensive self-esteem (i.e., high in explicit self-esteem and low in implicit self-esteem) had lower IAT scores if they had completed the experimental version of the conditioning task compared to their counterparts in the control condition. Further exploratory analyses revealed that this pattern of responding appeared primarily in men who began the study with defensive self-esteem. Indeed, when the data from women were analyzed alone a significant main effect of condition was found showing that overall women who completed the experimental version of the computer-based conditioning task had higher implicit self-esteem compared to women who completed the control version of the conditioning task. This finding will be addressed further in the General Discussion.

While there was an unexpected correlation between implicit and explicit self-esteem in Study 1, the lack of a correlation between implicit and explicit self-esteem across Studies 2, 3, and 4 provides evidence that implicit and explicit self-esteem represent two distinct dimensions of self-evaluation.

Replicating the finding from the previous studies, there were no effects found on explicit self-esteem. The lack of effects on explicit self-esteem across all four studies indicates that the conditioning task does not affect explicit self-evaluative responses, but rather affects the unconscious cognitive associations upon which implicit self-esteem is based. Furthermore, this finding supports a dual attitudes view of self-esteem. Although implicit self-esteem was influenced, no differences in explicit levels of self-esteem between conditions were reported.

CHAPTER 6

SUPPLEMENTARY ANALYSES

The primary goal of this dissertation was to test a cognitive interpersonal model of implicit self-esteem by using evaluative conditioning to enhance implicit self-esteem. Accordingly, the primary analyses presented in the previous chapters focus on that hypothesis. In this chapter I present supplementary analyses that explore the effects of the conditioning task in more detail, and on issues that are less central to the thesis. These include the results of a meta-analysis, findings from pooled data sets, and results from the Name Letter Task.

Also, in addition to enhancing implicit self-esteem through evaluative conditioning, I was interested in the effects of the conditioning task on people's reactions to socially provocative situations. To explore this, a measure of aggressive thoughts and feelings was included in Studies 2, 3, and 4. However, this topic is tangential to the focus of this thesis, thus the details and findings from the aggressiveness measure are reported in Appendix 3.

Overall Effect of the Conditioning Task on IAT Scores

To gain insight into the nature of the conditioning task on implicit self-esteem, I conducted a meta-analysis on the IAT main effect from Studies 1, 2, 3, and 4. Cohen's *d* was used as the measure of effect size. Comprehensive Meta-Analysis (1998) software was used to conduct the meta-analysis. The calculation of a composite *d* provided both a test of whether the value differed significantly from zero, and a 90% confidence interval

(CI). Because significance tests and CIs are largely redundant, I report only the CI here. I tested the effect of condition on post-measured implicit self-esteem scores, and the composite effect size was $d = 0.47$ (CI = 0.22, 0.70). This is a medium effect size (Cohen, 1988) and provides evidence that the conditioning task was having an overall positive influence on post-measured IAT scores compared to the control group.

To explore the role of pre-measured explicit self-esteem across all four studies, and to explore the data using regression analyses, I pooled the IAT data from Studies 1, 2, 3, and 4 and conducted a regression analysis on the IAT score. Results showed a main effect of condition, with no significant interaction with pre-measured explicit self-esteem, $\beta = .18$, $t(177) = 2.59$, $p = .01$. Participants who completed the experimental version of the conditioning task had a mean score of 223.52 on the IAT, while participants in the control condition had a mean score of 137.12. The effect remained significant when the baseline IAT value was excluded from the equation.

There were no significant effects of condition on post-measured explicit self-esteem or on mood. Furthermore, the significant main effect remained when mood and post-measured explicit self-esteem were included as covariates in the equation. There was no significant interaction between condition and gender; however, splitting the data file on gender revealed evidence of the main effect in both men and women.

This finding of only a strong main effect of condition in the pooled data set suggests that the two-way interaction in Study 1 was spurious. Furthermore, although the effect of pre-measured implicit self-esteem could not be assessed across all four studies (as the pre-measure of implicit self-esteem was not included in Studies 1 and 2), the results of the meta-analysis and of the regression analysis on the pooled data set suggest

that the three-way interaction in Study 4 may be tapping into the effects of individual differences in the sample (specifically, that defensive men may respond uniquely to the conditioning task). This is an interesting finding, and it will be addressed further in the General Discussion; however, the evidence suggests that on the whole the computer-based conditioning task enhances implicit self-esteem.

Name Letter Task

The Name Letter Task was included as both a pre- and post-measure of implicit self-esteem in Studies 3 and 4. As described earlier, in the Name Letter Task participants are asked to rate how much they "like" each letter of the alphabet. High rankings of one's own name letter initials indicate high implicit self-esteem. The results from the Name Letter Task were not as clear as those from the IAT. Furthermore, the IAT and post-measured Name Letter scores did not correlate with each other, indicating that they were tapping into different aspects of implicit self-esteem. As the IAT is, by nature, a measurement of cognitive associations, and because my hypothesis was based on the theory that implicit self-esteem is represented by cognitive self--*accepted* or self--*rejected* associations, I focused on the results from the IAT in the primary results.

The results from the Name Letter Task were not consistent across Studies 3 and 4. In order to obtain a clearer picture of the effect of the conditioning task on Name Letter scores, I merged the Name Letter data from Studies 3 and 4. The three-way interaction between condition, pre-measured implicit self-esteem and pre-measured explicit self-esteem was not significant. However, a significant two-way interaction between condition and pre-measured implicit self-esteem was found, $\beta = -.27$, $t(106) = -2.95$, $p =$

.004. Simple slopes analyses revealed that participants who began the study with low implicit self-esteem benefited the most from the computer-based conditioning task, $\beta =$ .28, $t(106) = 3.07$, $p = .003$. There was no significant effect of condition for participants who began the study high in implicit self-esteem. The two-way interaction between condition and pre-measured explicit self-esteem was not significant.

While, reminiscent of the 2-way interaction found in Study 1, it is unsurprising that participants who began the study low in implicit self-esteem benefited the most from the conditioning task, this finding may be a function of the nature of the Name Letter Task. The Name Letter Task contains a fixed scale ranging from 1 to 9, such that the highest rating participants can assign to any given letter is a 9. Descriptive statistics of the raw data of Name Letter scores showed that 48% of the participants in the experimental condition and 52% of the participants in the control condition rated their name letters as a 9. Thus, it is unsurprising that participants in the experimental condition who began the study with high implicit self-esteem did not differ from their counterparts in the control condition, as participants in both groups were already expressing the highest rating on the scale pre-manipulation. So, the two-way interaction may be an artifact of a ceiling effect of the Name Letter Task for participants who began the study with high implicit self-esteem. Overall, however, the findings on the Name Letter Task are encouraging in that they demonstrate that the conditioning task appears to have a positive effect on implicit self-esteem as measured by both the IAT and the Name Letter Task.

CHAPTER 7

GENERAL DISCUSSION

The four studies presented in this dissertation demonstrate that a computer-based conditioning task that pairs self-relevant information with photographs of accepting faces enhances implicit self-esteem. This finding provides support for the hypothesized interpersonal cognitive model of implicit self-esteem. While there was not perfect consistency across the four studies, a meta-analysis revealed that, overall, participants who repeatedly saw their self-relevant information followed by photographs of accepting faces had higher scores on the self-esteem IAT compared to participants who saw a random selection of accepting, rejecting, and neutral faces paired with their self-relevant information.

With regard to the individual studies, a main effect was found in Study 2 and Study 4; however, this was qualified by a two-way interaction between condition and pre-measured explicit self-esteem in Study 1, and by a three-way interaction between condition, pre-measured implicit self-esteem, and pre-measured explicit self-esteem in Study 4. Exploratory analyses revealed that defensive men did not appear to benefit from the positive effects of the conditioning task, and participants in the experimental condition generally were not aware of the contingency between their self-relevant information and photographs of smiling faces. No effects were found on post-measures of explicit self-esteem across all four studies. In the sections below, I discuss the theoretical implications of these findings.

A Cognitive Model of Implicit Self-Esteem

In the introduction, I proposed a cognitive model of implicit self-esteem that drew from an interpersonal model of self-esteem (Cooley, 1902; Baldwin & Sinclair, 1996; Leary, Tambor, et al., 1995) and a referential model of evaluative conditioning (De Houwer et al., 2001). Specifically, I hypothesized that implicit self-esteem is learned through evaluative conditioning whereby people learn cognitive associations between "self" and "accepted" or "self" and "rejected" through repeated exposure to interpersonal acceptance and rejection. To test this hypothesis, I developed a computer-based conditioning task that paired self-relevant information with photographs of smiling faces.

The hypothesis was supported by the results of the four studies presented in this dissertation. Pairing self-relevant information with social acceptance led to high scores on the self-esteem IAT.[8] No differences in mood were found between those participants who completed the experimental version of the computer-based conditioning task and those participants who completed a control version of the task, showing that pairing self-relevant information with photographs of smiling faces is unrelated to positive affect.

A Dual Attitudes Model of Self-Esteem

The results of the research presented in this dissertation strongly support a dual attitudes model of self-esteem. Creating cognitive self--*accepted* associations led to enhanced implicit self-esteem, yet no effects of the conditioning task were found on explicit self-esteem. This is very much in line with a dual attitudes model, which was

[8] Influencing implicit self-esteem through the evaluative conditioning procedure does not necessarily imply that implicit self-esteem can only be learned through repeated pairings between the self and social acceptance. However, the results from this body of research provide evidence to suggest that pairing the self with photographs of smiling faces leads to enhanced cognitive associations between "self" and "accepted."

developed in part to counter the argument that when a new attitude is formed, the new attitude replaces the old one in memory (Wilson et al., 2001). If this were the case with self-esteem, one would expect enhanced self-esteem at an implicit level to be reflected in enhanced reports of explicit self-esteem. This was not the case in the current research. Thus, people have the capacity to possess two different self-attitudes: one that is unconscious and implicit, and the other that is a conscious, explicit self-evaluation.

Considering the evidence supporting the existence of two distinct self-esteem systems, it is of interest to compare the interpersonal model of implicit self-esteem proposed in this dissertation to current interpersonal models of explicit self-esteem. To do this, I consider low self-esteem in particular. While both low implicit and low explicit self-esteem are related to social rejection, they differ in their cognitive representations. Explicit low self-esteem is related to the explicit perception that one is rejected or excluded by others (Heatherton & Wyland, 2003; Leary, Tambor, et al., 1995). For example, people with low explicit self-esteem (but not people with low implicit self-esteem) significantly rate their own anxiety as higher following an interview compared to people with high explicit self-esteem (Spalding & Hardin, 1999). This demonstrates that low explicit self-esteem is related to self-reported anxiety, while low implicit self-esteem is not. Explicit self-esteem involves the conscious and deliberative processing of social information related to the self. The end result is a conscious self-evaluation that reflects, for the most part, a person's beliefs and expectations about how he or she is regarded by others.

On the other hand, according to the cognitive model of implicit self-esteem supported by the research presented in this dissertation, people develop low implicit self-

esteem through repeated pairings of the self with interpersonal rejection, which creates an automatic self--*rejected* link. For example, people with low implicit self-esteem (but not people with low explicit self-esteem) were judged by an observer as more anxious following a self-relevant interview compared to people who were high in implicit self-esteem (Spalding & Hardin, 1999). Thus, while low implicit self-esteem is unrelated to explicit reports of anxiety, it is related to higher ratings of nonverbal anxiety as judged by another person. This suggests that implicit self-esteem is not only represented at an unconscious and automatic level, but is also expressed at an unconscious level. Cognitively, implicit self-esteem involves an automatic reaction to the self as a whole, where thoughts of the self or exposure to self-relevant information activate the self--*rejected* or self--*accepted* cognitive link. The end result is the "gut-level" (i.e., unconscious and automatic) attitude toward the self.

Recently some researchers have noted with concern the fact that measures of implicit self-esteem do not necessarily lead to the same conclusions as do measures of explicit self-esteem, and asked "How can we know which is really the better way to assess self-esteem?" (Heatherton & Wyland, 2003, p.227). Considering the evidence presented here for a dual attitudes model of self-esteem, the answer to this question may be that neither one is the better way to assess self-esteem; however, *both* implicit and explicit self-esteem should be measured in order to gain a complete picture of self-esteem.

Individual Differences in Implicit versus Explicit Self-Esteem

In Study 1 and Study 4, interactions between condition and pre-measured self-esteem were found. Specifically, Study 1 showed that participants who began the study

low in explicit self-esteem benefited most from the conditioning task, while Study 4 showed that participants who began the study with congruent high self-esteem (i.e., high in both implicit and explicit self-esteem) benefited most from the conditioning task. Furthermore, Study 4 showed that participants who began the study with a combination of high explicit/low implicit self-esteem showed significantly lower post-measured implicit self-esteem compared to their counterparts in the control condition. While the results from the meta-analysis as well as the pooled data showed a strong main effect of condition, exploratory analyses on the 3-way interaction revealed some interesting effects that are discussed below.

Gender. In an attempt to explain the three-way interaction from Study 4, I conducted some exploratory analyses involving gender. These exploratory analyses revealed that implicit self-esteem was enhanced for women in the experimental condition compared to women in the control condition. Furthermore, the three-way interaction in the full sample appeared specifically in men. The most puzzling finding was that men who began the study with defensive self-esteem (i.e., high explicit/low implicit) had lower scores on the IAT following the experimental version of the conditioning task compared to men who began the study with defensive self-esteem and completed the control version of the conditioning task.

Research in the area of explicit self-esteem has found that, in general, men gain self-esteem from success at tasks related to performance while women gain self-esteem from success in social situations (Josephs, Markus, & Tafarodi, 1992). It is possible that this same distinction exists in implicit self-esteem, which might explain why the effect of

the conditioning task was particularly evident in women in Study 4 (although analysis of the pooled data did show the main effect in both men and women).

Leary et al. (1995) found that women who were excluded rated themselves less positively than men, suggesting that women may be more sensitive to social exclusion than men. Furthermore, Pelham, Koole, Hardin, Hetts, Seah, and DeHart (2005) hypothesized that, due to gender socialization, women may be more attuned to their emotional states. They argue that because implicit self-associations, like other implicit attitudes, consist of more affective associations than do explicit attitudes, individuals who are more attuned with their emotions (i.e., women) would be more likely to show congruence between implicit and explicit self-esteem. Their findings showed that in general the correlation between implicit and explicit self-esteem is stronger in women than it is in men. I conducted exploratory analyses of the correlation found between implicit and explicit self-esteem in Study 1 and, in line with Pelham et al.'s (2005) results, found that the significant correlation appeared primarily in women. Given that women, in general, have been shown to have more congruent self-esteem than men, the results from Study 4 suggest that when men show incongruence, specifically in the combination of high explicit/low implicit, it may be representative of a cognitive mechanism that is distinct from self-esteem incongruence shown in women.

Cognitive mechanisms involved in congruent versus incongruent self-esteem. Some of the principles of the dual attitudes model might help explain the different cognitive mechanisms involved in congruent versus incongruent self-esteem. Wilson et al. (2000) outline different types of dual attitudes based on two aspects of automaticity: 1) awareness of the implicit attitudes; and 2) the capacity and motivation needed to override

the implicit attitude with the explicit attitude. When self-esteem is treated from the perspective of a dual attitudes model, researchers have generally implied an independent systems type of dual attitude. In an independent systems model, people are unaware of their implicit attitude and aware of their explicit attitude, and each independently develops evaluations. Notably, implicit attitudes influence implicit responses and explicit attitudes influence explicit responses.

Alternatively, people with defensive self-esteem might have their self-esteem represented by a type of dual attitude called *motivated overriding*. In motivated overriding, people are aware of their implicit attitude; however, they view this attitude as unwanted. They are motivated to override it with a more wanted attitude. Thus, people with defensive self-esteem might be aware that their implicit self-evaluations are negative (i.e., aware of a negative "gut feeling" of self-esteem), and view this as undesirable and unwanted. They are motivated to defend their explicit, positive self-evaluations in an attempt to prevent their implicit negative self-attitudes from being expressed. This view is supported by research showing that people with defensive self-esteem (i.e., high explicit/low implicit self-esteem) perceive favourable evaluations as more accurate than unfavourable evaluations (Bosson et al., 2002). At this point, and with the data available from the current research, it is difficult to theorize why men with defensive self-esteem showed lower scores on the IAT compared to men with defensive self-esteem in the control condition. However, future research should focus on determining whether implicit and explicit self-esteem operate according to different models of dual attitudes depending on whether or not self-esteem is congruent. Additionally, research should continue to explore the role of gender in the relationship between implicit and explicit self-esteem.

The Effects of Contingency Awareness

Results from Study 4 showed that participants in the experimental condition were not aware of the contingency between their self-relevant information and photographs of smiling faces. Exploratory analyses revealed that the effect remained even when participants who were aware of the contingency were excluded from the analysis. In the learning literature, there has been some discussion on the issue of contingency awareness in conditioning. As a general rule, contingency awareness is necessary for Pavlovian conditioning to occur (Fulcher and Hammerl, 2001a). If participants are not aware of the association between the conditioned stimulus (e.g., tone) and the unconditioned stimulus (e.g., food), learning will not occur. However, the role of contingency awareness in evaluative conditioning is less clear. Some researchers have argued that evaluative learning does, to some degree, depend on contingency awareness (see for example, Field, 2000). However, other researchers have argued that evaluative conditioning can and does occur without contingency awareness (Fulcher & Hammerl, 2001a). Furthermore, research has shown that when participants are aware of the contingency during evaluative conditioning they may attempt to discount the influence of the affective stimulus when evaluating the neutral stimulus (Fulcher & Hammerl, 2001b). Research has also shown that participants who are aware of the contingency sometimes show an effect in the opposite direction of what was expected (e.g., rating a neutral object negatively after it had been paired with a positive object).

Further exploratory analyses on the effects of contingency awareness suggested that participants in the experimental condition who were particularly aware of the contingency between presentation of the self-relevant information and photographs of

smiling faces may not have shown as strong of an effect of the conditioning task as those who were unaware of the contingency.

This finding once again speaks to the independence of implicit and explicit self-esteem and provides further support for a dual attitudes model, which theorizes that implicit attitudes are revealed on implicit measures while explicit attitudes are revealed on explicit measures (Wilson et al., 2000). Furthermore, and as stated by other researchers, this finding provides support for the argument that unconscious processes are not necessarily a "miniature" of conscious processes (Fulcher & Hammerl, 2001b). Participants in both the experimental and control conditions were not aware of the contingencies of the pairings between their self-relevant information and photographs of smiling faces (refer to Tables 1 and 2 for a review of the pairings between self-relevant information and face photographs in the experimental and control conditions). However, results show that the 100% pairing of self-relevant information with smiling faces in the experimental condition led to higher levels of implicit self-esteem compared to the control condition. Presumably, processing of the pairing was occurring at an unconscious level, the results of which were shown on measures of implicit self-esteem.

Recently I conducted a related study as a follow-up to the issue of contingency awareness in the conditioning of implicit self-esteem (Baccus & Baldwin, 2004). In this study, the computer-based conditioning task was modified to make the pairing between self-relevant information and photographs of smiling faces more obvious. Using only three simple line drawings to represent the smiling, frowning, and neutral faces in the computer-based conditioning task did this. Thus, participants saw the same smiling face paired with their self-relevant information on all 80 trials in the experimental condition.

Results showed that participants in the experimental group guessed that their self-relevant information was paired with the line drawing of the smiling face a significantly higher percentage of the time than participants in the control group (i.e., contingency aware). Furthermore, analyses of self-esteem measures showed that participants who began the study low in explicit self-esteem and subsequently completed the experimental version of the computer-based conditioning task had significantly higher post-measured explicit self-esteem than their counterparts in the control group. There were no significant differences between the groups on measures of implicit self-esteem (IAT and Name Letter Task). These findings once again speak to a dual attitudes model of self-esteem, and specifically suggest that implicit self-esteem is best addressed at an unconscious and automatic level, while explicit self-esteem is best addressed at a conscious and deliberative level. Thus, there is some evidence that contingency awareness decreases the effectiveness of the conditioning of implicit self-esteem, and perhaps enhances explicit self-esteem. However, more research is required.

Measurement of Implicit Self-Esteem

The self-esteem IAT is one of the most commonly used measures of implicit self-esteem and is presented in this dissertation as the primary measure of implicit self-esteem. While there is evidence supporting the validity of the IAT (Bosson et al., 2000), there are several questions as to what the self-esteem IAT is actually measuring. In the traditional calculation of the IAT score, as discussed earlier, reaction times to both self-related and other-related targets are included. A high score indicates that participants were faster to respond to self-related target words when "*Self*" and "*Good*" shared a

computer key and also that these same participants were faster to respond to other-related target words when *"Other"* and *"Bad"* shared a computer key. This is in contrast to the IAT calculation employed in the studies presented in this dissertation, in which only reaction times for self-related targets were used. Recently the meaningfulness of the role of the "other" in the self-esteem IAT has been questioned (Karpinski, 2004). Specifically, research has shown that changing the representation of the other—for example, by using the name of the participants' girlfriend or boyfriend as the target words for the "other" category—can change reaction times to self-related targets compared to unspecified other-related words (e.g., they, their) (Karpinski, 2004). Thus, assessing self-related targets separately from other-related targets appears to provide more detailed information regarding the cognitive processing that occurs during the IAT.

Furthermore, the self-esteem IAT administered to participants in the research presented in this dissertation was based on the standard format of the self-esteem IAT, in which the "good" and "bad" targets included words such as "sunshine" and "vomit." Although research has shown that target words such as these do not differ from a self-esteem IAT using positive trait words (Greenwald & Farnham, 2000), no research has investigated the effect of using words directly related to acceptance and rejection in the self-esteem IAT. For the studies in this dissertation, my main hypothesis involved a prediction that implicit self-esteem is conceptualized as cognitive associations between "self" and "accepted." However, rather than create a new IAT, I made the decision to use the standard and validated IAT. Further investigation of the conditioning of implicit self-esteem using the IAT as a dependent measure should consider testing a new version of the IAT with acceptance- and rejection- related words as the targets.

Greenwald and Farnham (2000) tested two versions of the self-esteem IAT during its development – one which assessed the associations of self versus other with good- and bad-related words (e.g., rainbow, vomit, sunshine, pain) and the other which assessed the associations of self versus other with positive and negative traits (e.g., smart, honest, ugly, weak). No significant differences were found between the two versions of the test. Currently, researchers tend to use the version of the self-esteem IAT where self- and other-related words are paired with good- and bad-related words (e.g., Bosson et al., 2000; Jordan et al., 2003).

Overall, there are nuances of the self-esteem IAT that remain unknown in implicit self-esteem research. In particular, future research should focus on understanding what is cognitively activated during the IAT. Specifically, research should concentrate on the role of each of the four target categories in the self-esteem IAT (i.e., self, other, good, bad) and explore how response times to each of these targets contribute to the cognitive representation of implicit self-esteem, and under what conditions their inclusion or exclusion in the total score matter.

The findings from the Name Letter Task were presented in Chapter 6. While the results were inconsistent across Studies 3 and 4, the merged data revealed that participants who began the study low in implicit self-esteem benefited the most from the conditioning task. The Name Letter Task was not correlated with the IAT, as shown in previous research (Bosson et al., 2000). Thus, it seems as though the IAT and the name Letter Task tap different aspects of implicit self-esteem. Arguably, the IAT is a more direct measure of cognitive association than the Name Letter Task, which assesses unconscious evaluations of self-related targets. The present studies did not allow for a

more thorough investigation of how the conditioning task influences implicit self-esteem as measured by the IAT compared to the Name Letter Task; however, future research should explore the distinction between the two measures.

Other Research on Changing Levels of Self-Esteem

As noted in an earlier footnote, while the research presented in this dissertation was being conducted (and reviewed for publication), researchers in other parts of the world were also exploring techniques to enhance self-esteem. At the University of Mannheim in Germany, Rikketa and Dauenheimer (2003) subliminally paired a self-referential (*I*) or non self-referential word (*Leo*) with either positive (*good*, *great*, and *valuable*) or negative trait words (*bad*, *lousy*, and *worthless*). The authors conclude that post-measured explicit self-esteem was enhanced for participants in the self-good pairing condition compared to participants in the self-bad condition. However, a review of the means within each block suggest that the key effect may have been a decrease in explicit self-esteem when *I* was paired with negative trait words, as opposed to an increase when *I* was paired with positive trait words. Furthermore, the authors found that pairing *Leo* with negative trait words had no effect on post-manipulation evaluations of Leo, which led them to conclude that their significant effects on self-esteem are more likely a result of the activation of evaluative self-knowledge as opposed to any changes in self-valence association. Finally, it is difficult to compare these findings to the findings presented in this dissertation, as Rikketa and Dauenheimer did not include any measures of implicit self-esteem.

At the University of Amsterdam, Dijksterhuis (2004) used subliminal evaluative conditioning in an attempt to increase implicit self-esteem. Participants were subliminally

presented with the Dutch word for "I" that was then followed by another subliminally presented word that in some cases was a positive trait word (e.g., *smart*, *beautiful*) and in others was a neutral word (e.g., *chair*, *bike*). Over the course of five studies, the main finding was that implicit self-esteem could be enhanced through the subliminal pairing of "I" with positive trait words. Upon closer examination, however, some key differences between Dijksterhuis's research and the research conducted in this dissertation become apparent.

The primary difference between the work of Dijksterhuis (2004) and the work presented here is the latter's focus on the interpersonal component of implicit self-esteem. I outlined in the introduction a theoretical model of implicit self-esteem that is strongly interpersonal. I specifically chose to pair self-relevant information with photographs of smiling faces to test the interpersonal aspect of the model. Dijksterhuis did not have this orientation, which becomes noticeable when the format of the Name Letter Task used by Dijksterhuis is explored.

A review of the literature has revealed that there are two different version of the Name Letter measure used in implicit self-esteem research: one which asks participants to rate the letters for how much they *like* each (Bosson et al., 2000), and the other which asks participants to rate the letter for how *attractive* or *beautiful* they find each letter (Koole et al., 2001). Dijksterhuis (2004) used the *attractive* wording on the Name Letter measure, while I used the *like* wording in the current research. Recently, the implications of these two different wordings have been explored (Sakellaropoulo & Baldwin, 2005). In two studies, participants were given both versions of the Name Letter measure along with various other dependent measures. Results found that the two version of the Name

Letter measure were only moderately correlated. Furthermore, an interesting interaction appeared whereby participants who scored low on the liking version of the Name Letter measure and high on the attractiveness version of the Name Letter Task showed high narcissism. Finally, priming participants with positive trait words tended to active narcissistic thoughts and feelings. Sakellaropoulo & Baldwin (2005) suggest that, like explicit self-esteem, implicit self-esteem should not be viewed as a one-dimensional construct.

In Dijksterhuis's research, and based on the findings of Sakellaropoulo and Baldwin (2005) it appears as though subliminally pairing "I" with positive trait words may influence self-admiration as opposed to self-acceptance. Dijksterhuis found that pairing "I" with positive trait words enhanced scores on the "attractiveness" version of Name Letter Task. Sakellaropoulo and Baldwin found that both positive trait words and high scores on the "attractiveness" version of the Name Letter Task are related to narcissism. Future research should explore the possible two-dimensional nature of implicit self-esteem, and how different types of evaluative conditioning can affect each dimension.

Overall, while the research of Riketta and Dauenheimer (2003) and Dijksterhuis (2004) are generally in line with the application of conditioning to self-esteem, the research presented in this dissertation specifically addresses the interpersonal component of implicit self-esteem, which concurrent research on the conditioning of self-esteem has ignored.

Limitations and Directions for Future Research

All four studies presented in this dissertation were conducted using a student population sample of men and women, which limits the generalizability of the findings. There are some limitations when it comes to assessing implicit self-esteem in a younger population. The IAT is not traditionally used in populations of children and adolescents. Indeed, the task is cognitively demanding and children may find it difficult to do, which would detract from the sensitivity of the measure. Furthermore, as the development of implicit self-esteem has not been decisively mapped out, it is unknown if and how implicit self-esteem is expressed in children and adolescents. Future research should focus on developing a measure of implicit self-esteem for use in children and adolescents, and explore the effects of conditioning with this age group.

It is unknown how long the effects of the conditioning will last. In the four studies presented here, participants' implicit self-esteem was measured directly following the computer-based conditioning task. Theoretically, it is unlikely that exposure to a 5-minute computer conditioning task will lead to permanent changes to implicit self-esteem. However, according to the principles of learning theory, the repeated exposure to the pairing of "self" with "acceptance" should lead to a stronger cognitive self--*accepted* association. Future research should focus on assessing the long-term effects of the conditioning task. Although the long-term effects of the conditioning task were not tested in the current research, this should not overshadow the importance of the finding that implicit self-esteem can be influenced by a five-minute computer task.

Conclusion

The construct of implicit self-esteem is emerging as an integral component in the understanding of self-esteem as a whole. In the introduction of this dissertation, I pointed out that there was a gap in current research regarding implicit self-esteem. Specifically, I noted that little was known regarding the cognitive representation of implicit self-esteem, how implicit self-esteem is learned, and how implicit self-esteem can be changed. The theory and research presented in this dissertation have provided original empirical contributions to all three of these areas.

I have proposed a theory regarding the cognitive representation of implicit self-esteem. Drawing from research supporting the interpersonal nature of self-esteem (Leary, Tambor, et al., 1995; Baldwin & Sinclair, 1996), and theories of evaluative conditioning (De Houwer et al., 2001), I put forth an interpersonal cognitive model of implicit self-esteem stating that implicit self-esteem is cognitively represented as associations between self--*accepted* or self--*rejected*. The research presented in this dissertation supports this interpersonal cognitive model of implicit self-esteem and reveals that implicit self-esteem can be enhanced through evaluative conditioning—i.e., by pairing self-relevant information to interpersonal acceptance.

REFERENCES

Aiken, L.S., & West. S.G. (1991). *Multiple regression: Testing and interpreting interactions*. London: Sage.

Baccus, J.R., Baldwin, M.W., & Packer, D.J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science, 15*, 498-502.

Baccus, J.R., & Baldwin, M.W. (2004). *The effect of contingency awareness on the conditioning of implicit self-esteem.* (Unpublished data).

Baeyens, F., Eelen, P., & Van Den Bergh, O. (1990). Contingency awareness in evaluative conditioning: A case for unaware affective-evaluative learning. *Cognition and Emotion, 4*, 3-18.

Baeyens, F., Eelen, P., Van Den Bergh, O., & Crombez, G. (1989). Acquired affective-evaluative value: Conservative but not unchangeable. *Behaviour Research and Therapy, 27*, 279-287.

Baldwin, M.W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin, 112*, 461-484.

Baldwin, M.W., & Main, K.J. (2001). Social anxiety and the cued activation of relational knowledge. *Personality and Social Psychology Bulletin, 27*, 1637-1647.

Baldwin, M.W., & Sinclair, L. (1996). Self-esteem and "if...then" contingencies of interpersonal acceptance. *Journal of Personality and Social Psychology, 71*, 1130-1141.

Baumeister, R.F. (1993). *Self-esteem: The puzzle of low self-regard.* New York, NY: Plenum Press.

Bosson, J.K., Brown, R.P., Zeigler-Hill, V., & Swann, W.B. (2003). Self-enhancement tendencies among people with high explicit self-esteem: The moderating role of implicit self-esteem. *Self and Identity, 2,* 169-187.

Bosson, J.K., Swann, W.B., & Pennebaker, J.W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology*, 79, 631-643.

Brewin, C. (1989). Cognitive change processes in psychotherapy. *Psychological Review*, *96*, 379-394.

Brown, J.D. (1991). Accuracy and bias in self-knowledge. In C.R. Snyder and D.F. Forsyth (Eds.), *Handbook of Social and Clinical Psychology: The Health Perspective*. New York: Pergamon Press, pp. 158-178.

Brown, J.D., & Dutton, K.A. (1995). The thrill of victory, the complexity of defeat: Self-esteem and people's emotional reactions to success and failure. *Journal of Personality and Social Psychology, 68,* 712-722.

Brown, R.P., & Bosson, J.K. (2001). Narcissus meets Sisyphus: Self-love, self-loathing, and the never-ending pursuit of self-worth. *Psychological Inquiry, 12,* 210-213.

Buckley, K.E., Winkel, R.E., & Leary, M.R. (2004). Reactions to acceptance and rejection: Effects of level and sequence of relational evaluation. *Journal of Experimental Social Psychology*, *40*, 14-28.

Cacioppo, J.T., Marshall-Goodell, B.S., Tassinary, L.G., & Petty, R.E. (1992). Rudimentary determinants of attitudes: Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high. *Journal of Experimental Social Psychology, 28, 207*-233.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences,* 2nd Ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Comprehensive Meta-Analysis (1998). Biostat.

Cooley, C.H. (1902). *Human nature and social order.* New York: Charles Scribner & Sons.

Cunningham, W.A., Preacher, K.J., & Banaji, M.R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12,* 163-170.

Dasgupta, N., & Greenwald, A.G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81,* 800-814.

Dawson, M.E., & Schell, A.M. (1987). Human autonomic and skeletal classical conditioning: The role of conscious cognitive factors. In G. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans* (pp. 27-56). Chichester, England: Wiley.

De Houwer, J. (2001). Contingency awareness and evaluative conditioning: When will it be enough? *Consciousness and Cognition, 10,* 550-558.

De Houwer, J., Baeyens, F., & Hendrickx, H. (1997). Implicit learning of evaluative associations. *Psychologica Belgica, 37,* 115-130.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127,* 853-869.

Diener, E., & Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *Journal of Personality and Social Psychology, 68,* 653-663.

Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology, 86,* 345-355.

Dodgson, P.G., & Wood, J.V. (1998). Self-esteem and the cognitive accessibility of strengths and weaknesses after failure. *Journal of Personality and Social Psychology, 75,* 178-197.

Domjan, M. (2000). *The Essentials of Conditioning and Learning, Second Edition.* Nelson/Thomson Learning, Scarborough, Ontario.

Dovidio, J.F., & Gaertner, S.L. (1998). On the nature of contemporary prejudice: The causes, consequences, and challenges of aversive racism. In J.L. Eberhardt, & S.T. Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 3-32). Thousand Oaks, CA: Sage Publications, Inc.

Dutton, K.A., & Brown, J.D. (1997). Global self-esteem and specific self-views as determinants of people's reactions to success and failure. *Journal of Personality and Social Psychology, 73,* 139-148.

E-Prime (2002). (Version 1.0): Psychology Software Tools Inc.

Farnham, S.D., Greenwald, A.G., & Banaji, M.R. (1999). Implicit self-esteem. In D. Abrams & M.A. Hogg (Eds.), *Social identity and social cognition* (pp. 230-248). Oxford, England: Blackwell Publishers.

Field, A.P. (2000). I like it but I'm not sure why: Can evaluative conditioning occur without conscious awareness? *Consciousness and Cognition, 9*, 13-36.

Fulcher, E.P. & Hammerl, M. (2001a). When all is revealed: A dissociation between evaluative learning and contingency awareness. *Consciousness and Cognition, 10*, 524-549.

Fulcher, E.P. & Hammerl, M. (2001b). When all is considered: Evaluative learning doe not require contingency awareness. *Consciousness and Cognition, 10*, 567-573.

Greenwald, A.G., & Banaji, M.R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4-27.

Greenwald, A.G., & Farnham, S. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79, 1022-1038.

Greenwald, A.G., Banaji, M.R., Rudman, L.A., Farnham, S.D., Nosek, B.A., & Mellott, D.S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3-25.

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464-1480.

Hammerl, M., Bloch, M., & Silverthorne, C.P. (1997). Effects of US-alone presentations on human evaluative conditioning. *Learning and Motivation, 28*, 191-194.

Harter, S. (1993). Causes and consequences of low self-esteem in children and adolescents. In R.F. Baumeister (Ed.), *Self-Esteem: The Puzzle of Low Self-Regard*, Plenum Press, New York.

Heatherton, T.F., & Polivy, J. (1991). Development and validation of a scale for measure state self-esteem. *Journal of Personality and Social Psychology, 60*, 895-910.

Heatherton, T.F., & Wyland, C.L. (2003). Assessing self-esteem. In S.J. Lopez & C.R. Synder (Eds.), *Positive Psychology Assessment* (pp. 219-233).

Johnsrude, I.S., Owen, A.M., Zhao, W.V., & White, N.M. (1999). Conditioned preferences in humans: A novel experimental approach. *Learning and Motivation, 30*, 250-264.

Jones, J.T., Pelham, B.W., & Mirenberg, M.C. (2002). Name letter preferences are not merely mere exposure: Implicit egotism as self-regulation. *Journal of Experimental Social Psychology, 38*, 170-177.

Jordan, C.H., Spencer, S.J., & Zanna, M.P. (2002). "I love me...I love me not": Implicit self-esteem, explicit self-esteem, and defensiveness. In S.J. Spencer, S. Fein, M.P. Zanna, & J. M. Olson (Eds.), *The Ontario Symposium Vol. 9. Motivated social perception* (pp. 117-145).

Jordan, C.H., Spencer, S.J., Zanna, M.P., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive self-esteem. *Journal of Personality and Social Psychology, 85*, 969-978.

Josephs, R.A., Markus, J.R., & Tafarodi, R.W. (1992). Gender and self-esteem. *Journal of Personality and Social Psychology, 63,* 391-402.

Karpinski, A. (2004). Measuring self-esteem using the implicit association test: The role of other. *Personality and Social Psychology Bulletin, 30*, 22-34.

Karpinski, A. & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of Personality and Social Psychology, 81*, 774-788.

Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry, 14*, 1-26.

Kitayama, S., & Karasawa, M. (1997). Implicit self-esteem in Japan: Name letters and birthday numbers. *Personality and Social Psychology Bulletin, 23*, 736-742.

Koole, S.L., Dechesne, M., & van Knippenberg, A. (2000). *The sting of death: Evidence that reminders of mortality undermine implicit self-esteem.* Unpublished manuscript.

Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology, 80*, 669-685.

Koole, S.L., & Pelham, B.W. (2003). On the nature of implicit self-esteem: The case of the name letter effect. In S.J. Spencer, S. Fein, M.P. Zanna, & J. M. Olson (Eds.), *The Ontario Symposium Vol. 9. Motivated social perception* (pp. 93-116).

Koole, S.L., Smeets, K., van Knippenberg, A., & Dijksterhuis, A. (1999). The cessation of rumination through self-affirmation. *Journal of Personality and Social Psychology, 77*, 111-125.

Leary, M.R., & Baumeister, R.F. (2000). The nature and function of self-esteem: Sociometer theory. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 127-155). Hillsdale, NJ: Lawrence Erlbaum Associates.

Leary, M.R., Schreindorfer, L.S., & Haput, A.L. (1995). The role of low self-esteem in emotional and behavioral problems: Why is low self-esteem dysfunctional? *Journal of Social and Clinical Psychology, 14*, 297-314.

Leary, M.R., Tambor, E.S. Terdal, S.K., & Downs, D.L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social Psychology*, 68, 518-530.

Lieberman, D.A. (1993). *Learning, Behavior, and Cognition. Second Edition.* Brooks/Cole, Pacific Grove, California.

Lorr, M., & McNair, D. (1982). The profile of mood states inventory. San Diego, CA: Educational and Industrial Testing Service.

MacDonald, F., Saltzman, J.L., & Leary, M.R. (2003). Social approval and trait self-esteem. *Journal of Research in Personality, 37*, 23-40.

Mead, G.H. (1934). *Mind, self, and society.* Chicago: University of Chicago Press.

Nuttin, J.M., Jr. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology, 15*, 353-351.

Nuttin, J.M., Jr. (1987). Affective consequences of mere ownership: The name letter effect. *European Journal of Social Psychology, 17*, 381-402.

Olson, M.A., & Fazio, R.H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science, 12*, 413-417.

Olson, M.A., & Fazio, R.H. (2002). Implicit acquisition and manifestation of classically conditioned attitudes. *Social Cognition, 20*, 89-104.

Paulhus, D.L. (1986). Self-deception and impression management in test responses. In A. Angleitner and J.S. Wiggins (Eds.), *Personality Assessment via Questionnaire*. New York: Springer-Verlag, pp. 143-165.

Pavlov, I.P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford University Press, London.

Pelham, B.W., Koole, S.L., Hardin, C.D., Hetts, J.J., Seah, E., & DeHart, T. (2005). Gender moderates the relation between implicit and explicit self-esteem. *Journal of Experimental Social Psychology*, 41, 84-89.

Rescorla, R.A. (1988) Pavlovian conditioning: It's not what you think it is. *American Psychologist, 43,* 151-160.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64-99). New York: Appleton-Century-Crofts.

Riketta, M., & Dauenheimer, D. (2003). Manipulating self-esteem with subliminally presented words. *European Journal of Social Psychology, 33*, 679-699.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Sakellaropoulo, M., & Baldwin, M. (2005). *The hidden sides of self-esteem: Two dimensions of implicit self-esteem and their relation to narcissistic reactions.* Manuscript submitted.

Schimmack, U., & Diener, E. (2003). Predictive validity of explicit and implicit self-esteem for subjective well-being. *Journal of Research in Personality, 37*, 100-106.

Shimizu, M., & Pelham, B.W. (2004). The unconscious cost of good fortune: Implicit and explicit self-esteem, positive life events, and health. *Health Psychology, 23*, 101-105.

Spalding, L.R., & Hardin, C.D. (1999) Unconscious unease and self-handicapping: Behavioral consequences of individual differences in implicit and explicit self-esteem. *Psychological Science, 10*, 535-539.

Staats, A.W., & Staats, C.K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal Psychology, 57*, 37-40.

Tafarodi, R.W., & Swann, W.B. Jr. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of a measure. *Journal of Personality Assessment, 65,* 322-342.

Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology, 82*, 919-934.

Wilson, T.D., Lindsey, S., & Schooler, T.Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101-126.

Zanna, M.P. Kiesler, C.A., & Pilkonis, P.A. (1970). Positive and negative attitudinal affect established by classical conditioning. *Journal of Personality and Social Psychology, 14*, 321-328.

Appendix 1: Depiction of the Computer-Based Conditioning Task

In the computer-based conditioning task, participants use the computer mouse to click on target words when they appear in one of four quadrants on the screen. In the experimental version of the conditioning task (Figure A), self-relevant targets (e.g., first name, birthday) are always followed by a photograph of a smiling face. In this example, "Jodene" and "June 21" (panels 1 and 5) are self-relevant pieces of information, and thus are followed by a smiling face both times (panels 2 and 6). A photograph of either a smiling or a neutral face follows non-self-relevant targets.

In the control version of the self-esteem conditioning task (Figure B), participants also use the computer mouse to click on target words when they appear in one of four quadrants on the screen. However, in the control version of the conditioning task, both self-relevant and non-self-relevant targets are followed by a random selection of smiling, frowning, and neutral faces. In this example, and compared to Figure 1, the self-relevant pieces of information (panels 1 and 5) are now followed by a neutral face (panel 2) and a smile (panel 6).

Figure A: Experimental Condition

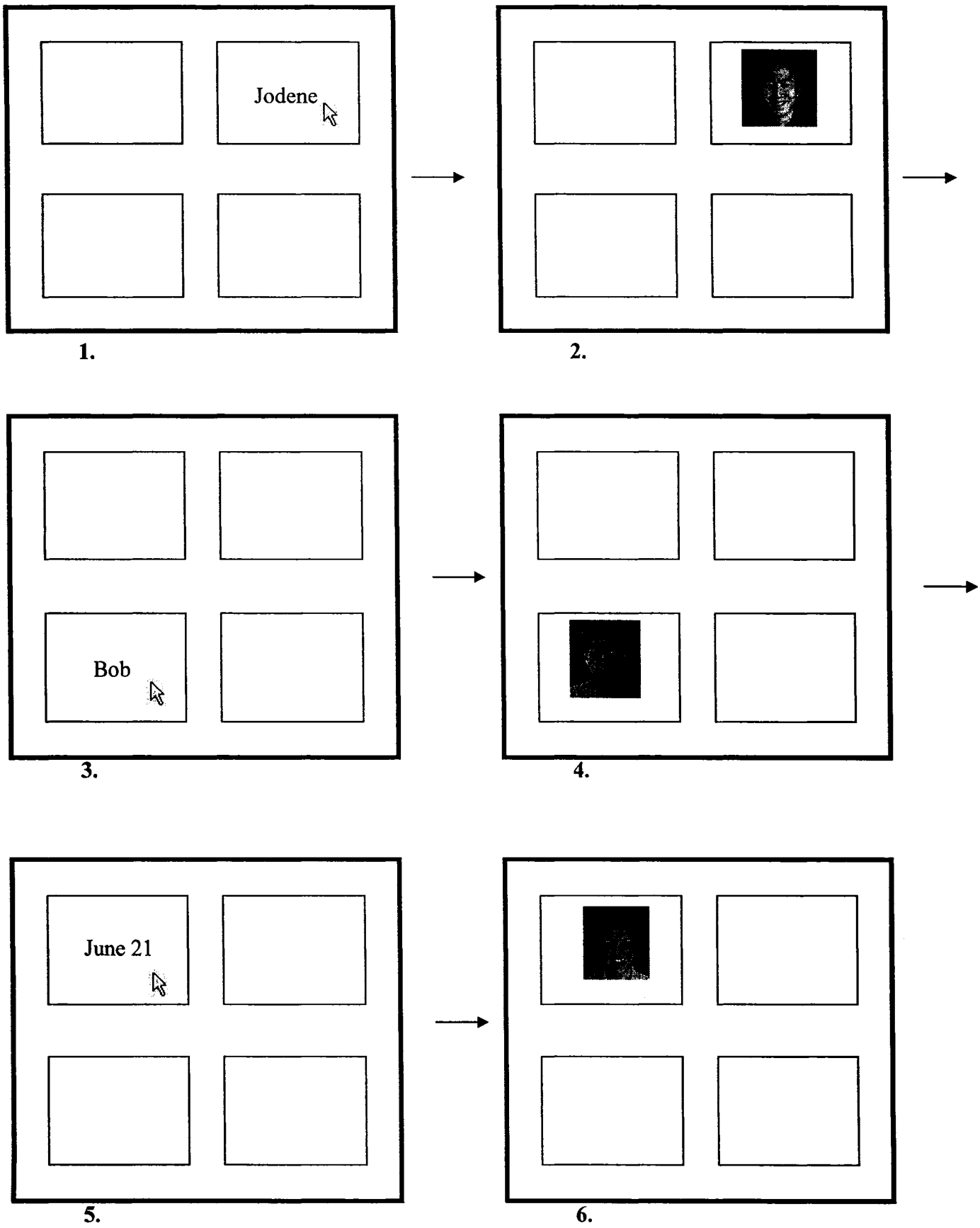
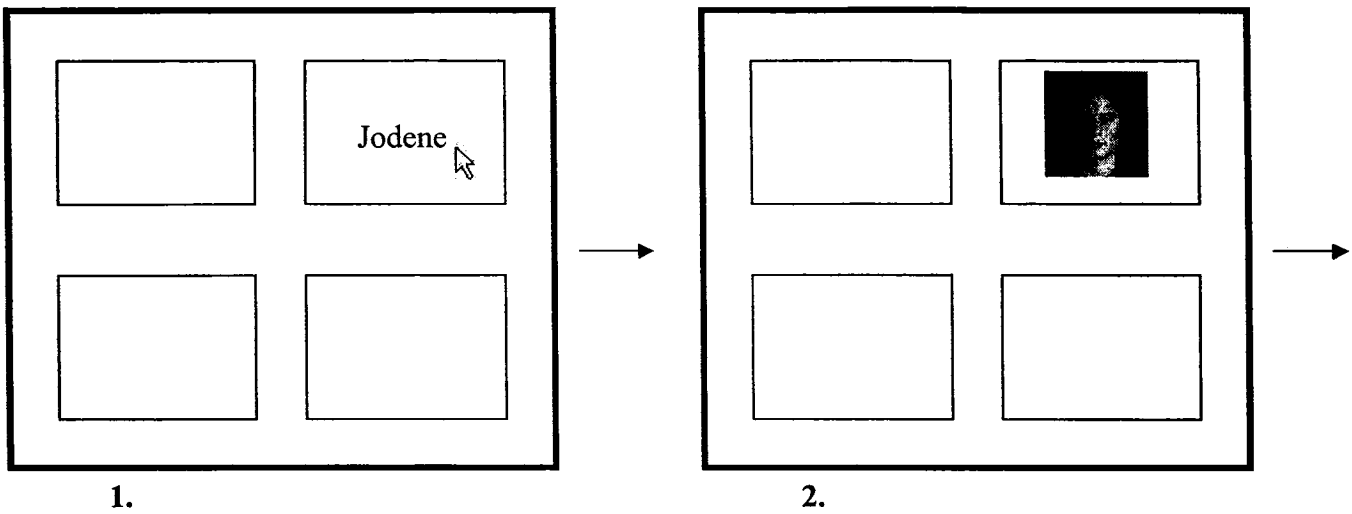
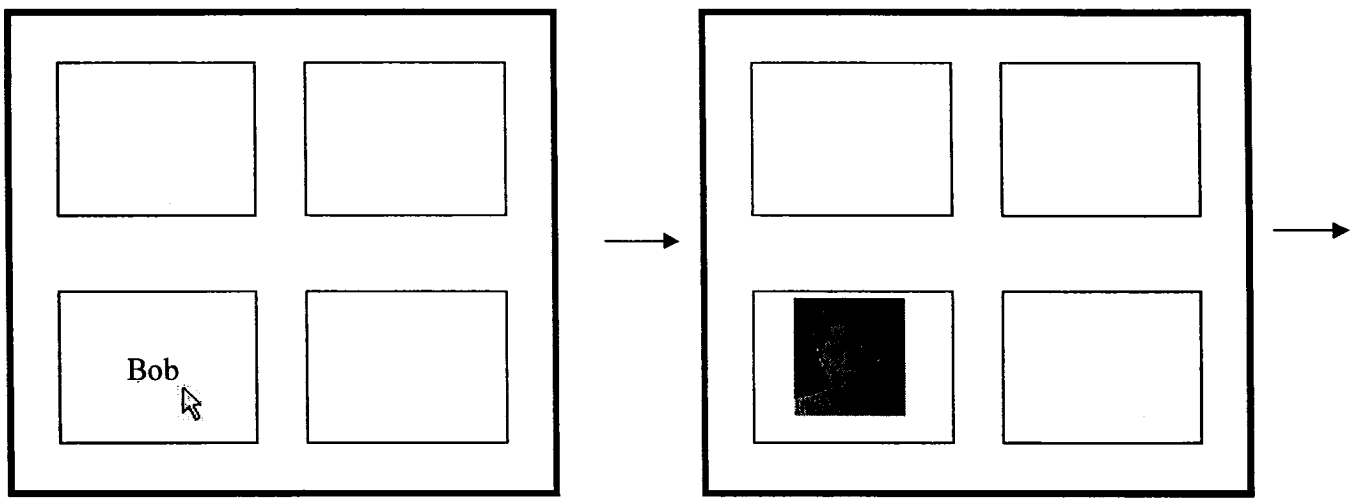


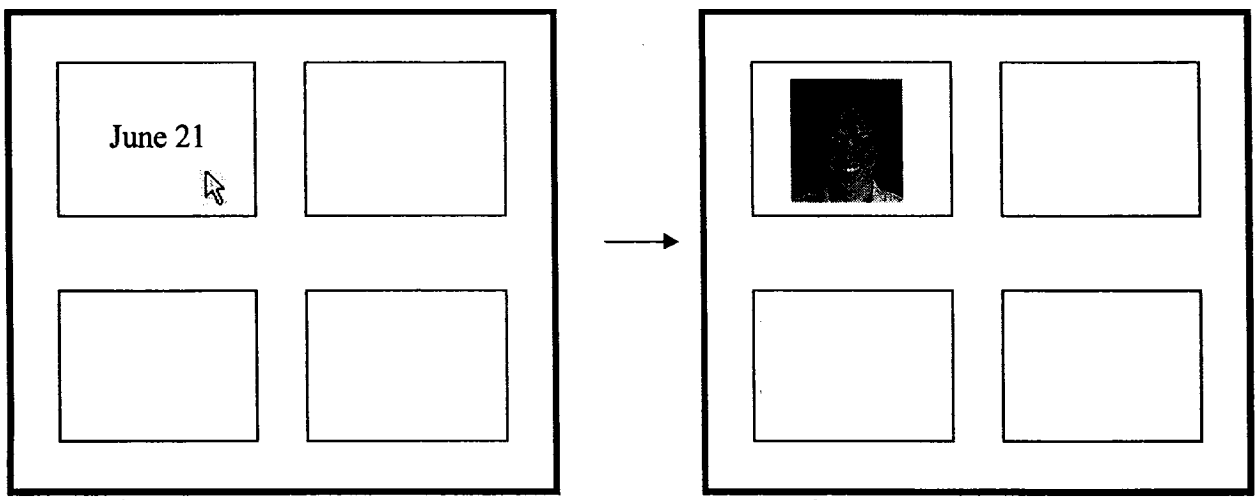
Figure B: Control Condition



1. 2.



3. 4.



5. 6.

Appendix 2: Words used in the self-esteem IAT

Self	Other	Good	Bad
I	they	rainbow	pain
me	them	happy	death
my	their	smile	poison
mine	it	joy	grief
self	other	warmth	agony
myself	themselves	pleasure	sickness
		paradise	tragedy
		sunshine	vomit

Appendix 3:

Aggressive Thoughts and Feelings in Relation to the Conditioning Task

The computer-based conditioning task was primarily developed as a means to test the hypothesized cognitive model of implicit self-esteem. However, I also wanted to assess the impact of the conditioning task on social behaviour that might be related to implicit self-esteem: aggressiveness. Research on the link between aggression and explicit self-esteem has been confusing, with some researchers suggesting that narcissism (i.e., an inflated sense of self) might contribute to aggression more so than low self-esteem (Bushman & Baumeister, 1998) and others suggesting that a low self-concept can trigger aggressive troublemaking behaviour (March, Parada, Yeung, & Healey, 2001). Furthermore, recent research has highlighted a link between violent video games and increased aggression (Anderson & Dill, 2000). To date, the link between implicit self-esteem and aggression has not been explored.

To assess participants' aggressive thoughts and feelings following the conditioning task, Taylor's (1967) behavioural aggression paradigm was adapted into written form (included at the end of the section). Participants were given three scenarios; in each they were asked to imagine competing against another student (who was seated in a different room) on a computer game. In two of the scenarios, the participant was asked to imagine that the other student had previously insulted or rejected him or her; the third scenario was a neutral version where no previous interaction with the other student had occurred. Participants read that they had won the computer competition and now had the opportunity to blast the other student with some noise. They were asked to indicate how

loud and how long they would like to set the noise blast for. The three scenarios were counterbalanced across participants. This measure was included in Studies 2, 3, and 4.

The results regarding aggression across the Studies 2, 3, and 4 were not clear, and in fact no significant effects were found in Study 3. Furthermore, aggressiveness scores were unrelated to IAT scores. That is, although participants who completed the experimental version of the task had higher IAT scores than those who completed the control version of the task, these increased IAT scores did not contribute to the relationship between condition and aggressiveness. This suggests that the low aggressiveness scores in the experimental condition were not a result of enhanced implicit self-esteem per se, but rather that conditioning self-relevant information to social acceptance leads more directly to less aggressive thoughts and feelings compared to participants who did not receive the conditioning. However, there were interesting findings in Study 2 and Study 4, and further analyses on the pooled data set revealed additional significant findings.

Aggressiveness – Study 2

A regression analysis on a total aggressiveness score (mean score across all three scenarios) did not result in any significant effects. I then analyzed each scenario separately. A regression analysis on the aggressiveness score from the rejection scenario with condition, pre-measured explicit self-esteem, and the cross-product interaction between the two as predictors showed no significant interaction between condition and explicit self-esteem. The interaction term was dropped from the equation.

Condition had a marginally significant effect on aggressive thoughts and feelings, $\beta = -.42$, $t(19) = -2.00$. $p = .060$, indicating that participants who completed the

experimental version of the conditioning task had less aggressive thoughts and feelings than those who completed the control version of the task. While participants in the control condition showed a mean aggression score of 3.41, those in the experimental condition had a mean of 2.20.

Aggressiveness-Study 4

The aggressiveness score was uncorrelated with the self-target IAT score, r = .056, *ns*. A regression analysis revealed a significant two-way interaction between condition and pre-measured explicit self-esteem was found, β = .54, $t(53) = 3.05$, p = .004. Simple slopes analyses revealed that participants who began the study low in explicit self-esteem had significantly lower aggressiveness scores if they had been in the experimental condition, compared to those participants with low explicit self-esteem in the control condition, β = -.46, $t(53) = -2.69$, p = .01. Participants who began the study with low explicit self-esteem and completed the experimental version of the conditioning task had a mean aggressiveness score of 3.06, compared to a mean of 3.84 for their counterparts in the control condition. There was no significant difference in aggressiveness score between the experimental and control condition for those participants who began the study high in explicit self-esteem.

Aggressiveness – Pooled Data Set

Using the pooled data set presented earlier for the exploration of the IAT findings, I regressed condition, pre-measured explicit self-esteem and the interaction between the two on the total aggressiveness scores from all three studies. Results showed a significant two-way interaction between condition and pre-measured explicit self-esteem, β = .34, $t(127) = 2.86$, p = .005. Analysis of the simple slopes showed that participants in the

experimental condition who began the study with low explicit self-esteem showed lower aggressiveness scores compared to their counterparts in the control condition, $\beta = -.34$, $t(127) = -2.84$, $p = .005$. There were no differences between conditions for participants who began the study high in explicit self-esteem.

Aggressiveness and the Conditioning of Implicit Self-Esteem

Overall, the aggression findings highlight the relationship between interpersonal acceptance, low explicit self-esteem, and aggression. While some research has theorized that narcissism (which is related to high explicit self-esteem) plays a larger contributing role to aggressiveness than does low explicit self-esteem (Bushman & Baumeister, 1998), the findings presented here show that low explicit self-esteem is related to aggressiveness. Furthermore, exposure to a computer-based conditioning task that creates a sense of social acceptance leads to lower aggressive thoughts and feelings for people with low explicit self-esteem.

Aggressiveness Measure

Please read the following situation and answer the questions below:

You are taking part in a psychology experiment. You enter into the designated room, where another student is already waiting.

(The Rejected and Insult paragraphs were inserted here to create the other two scenarios)

The experimenter sends you each to a different room, and tells you that you will be playing a video game against each other. The two computers are connected to each other, and the goal of the game is for you to push a button on your computer faster than your opponent does on their computer. The winner of each trial gets to "punish" the other player with a blast of noise. For this blast of noise, you are allowed to set the noise level, and to set the duration of the noise. You win the first trial.

1. What level do you FEEL like setting the noise at? (please circle)

Very Low	*Low*	*Med. Low*	*Mediuum*	*Med. High*	*High*	*Very High*
barely noticeable	annoying	very annoying	disturbing	somewhat painful	quite painful	extreme pain

2. What level do you ACTUALLY set the noise at? (please circle)

Very Low	*Low*	*Med. Low*	*Mediuum*	*Med. High*	*High*	*Very High*
barely noticeable	annoying	very annoying	disturbing	somewhat painful	quite painful	extreme pain

3. How long do you FEEL like setting the duration of the noise (in seconds)?

1 2 3 4 5

4. How long do you ACTUALLY set the duration of the noise (in seconds)?

1 2 3 4 5

(Rejection Paragraph)

The experimenter tells you that the first part of the study involves filling out several personality questionnaires, and she leads each of you to a separate room to do so. After you finish filling them out, the experimenter takes the questionnaires from you, and tells you that the next part of the study gives you the opportunity to work with the other person. Before that, each student will read through the completed questionnaires to get an idea of each other's personality. She leaves the room with your questionnaire and returns awhile later to say that the student has decided, after reading your completed questionnaires, that they would rather complete the second part of the study without you. You finish the second part of the study, some word puzzles, by yourself.

(Insult Paragraph)

The experimenter tells you that the first part of the study is a vocabulary test. For each of you, she will hold up various words and you are to say them out loud as quickly as possible. The other student goes first, and completes the task quickly. When it is your turn, you encounter some difficult words that you are not sure how to pronounce. As you stumble over some of the words, the other student snickers at you and says "at least some of us got past a grade 6 reading level".

REFERENCES

Anderson, C.A., & Dill, K.E. (2000). Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *Journal of Personality and Social Psychology, 78*, 772-790.

Bushman, B.J., & Baumeister, R.F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology, 75*, 219-229.

Marsh, H.W., Parada, R.H., Yeung, A.H., & Healey, J. Aggressive school troublemakers and victims: A longitudinal model examining the pivotal role of self-concept. *Journal of Educational Psychology, 93*, 411-419.

Taylor, S.P. (1967). Aggressive behavior and psychological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality, 35*, 297-310.

McGill University

ETHICS REVIEW
ANNUAL STATUS REPORT/RENEWAL REQUEST/FINAL REPORT

Continuing review of human subjects research requires, at a minimum, the submission of an annual status report to the REB. This form must be completed to request renewal of ethics approval. If a renewal is not received before the expiry date, the project is considered no longer approved and no further research activity may be conducted. When a project has been completed, this form can also be used as a Final Report, which is required to properly close a file. To avoid expired approvals and, in the case of funded projects, the freezing of funds, this form should be returned at least **1 month** before the current approval expires.

REB File #: REB#49-0901
Project Title: Shifting levels of implicit self-esteem via classical conditioning II
Principal Investigator: Mark Baldwin
Department/Phone/Email: Psychology/398-6090/mbaldwin@ego.psych.mcgill.ca
Faculty Supervisor (for student PI):

1. Were there any significant changes made to this research project that have any ethical implications? ___Yes __X__No
   If yes, describe these changes and append any relevant documents that have been revised.

2. Are there any ethical concerns that arose during the course of this research? ___ Yes __X_ No. If yes, please describe.

3. Have any subjects experienced any adverse events in connection with this research project? ___ Yes __X_ No
   If yes, please describe.

4. __X__ This is a request for renewal of ethics approval.

5. ____ This project is no longer active and ethics approval is no longer required.

6. List all current funding sources for this project and the corresponding project titles **if not exactly the same** as the project title above. Indicate the Principal Investigator of the award if not yourself.

SSHRC #410-2002-0646. Self-esteem and relational schemas: Principles of Change

Principal Investigator Signature: _____ Date: Oct. 4/04

Faculty Supervisor Signature: _____ Date: _____
(for student PI)

Submit to Lynda McNeil, Research Ethics Officer, James Administration Bldg., rm 429, fax: 398-4853

dministrative Use **REB:** ___AGR ; ___EDU ___REB-I ___REB-II

__ The closing report of this terminated project has been reviewed and accepted

__ The continuing review for this project has been reviewed and approved

__ Expedited Review ___Full Review

gnature of REB Chair or designate: _____ Date: Oct 6, 2004

pproval Period: Nov 1, 2004 to Oct 31 2005

Submit to Lynda McNeil, Research Ethics Officer, James Administration Bldg., rm 429, fax: 398-4853

(version October 2002)

Done Back Save Help

49-0901 - PROF Mark Baldwin
10SOCIAL SCIENCES & HUMANITIES RESEARCH COUNCIL
Shifting Levels of Implicit Self-Esteem Via Classical Conditioning II

Management

Summary

Institution Number 49-0901 **Change**

Submission Date 21-Oct-2003

Original Meeting Date

Board Name University Research Ethics Board II **Set**

Reviewers None Selected **Set**

Approval

Status Approved

Approved From 01-Nov-2004

Approved To 31-Oct-2005

Review Category Expedited

Short Title

Full Title Shifting Levels of Implicit Self-Estee
Classical Conditioning II

Objectives and Purposes

Overview

User-Defined Fields

PI Status Faculty

Student PI Name

Other Researchers/Status Jodene Baccus- co-PI
Genevieve Taylor

Research Type Faculty Research

Hospital Site DH
JGH
MUHC-MCH