STEREOTYPE THREAT IMPAIRS THE FEELING OF LEARNING

by

William Hall

B.A., University of Waterloo, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2012

© William Hall, 2012
Abstract

The present study examined stereotype threat impairments on stigmatized learners’ ability to develop conscious awareness of what they have learned, even when learning has taken place. To test this, participants completed a task where learning is initially implicit but the “feeling of learning” develops with greater experience. Participants were female undergraduates who completed an implicit category learning task under threat or control conditions. Across 192 trials, participants made a category choice, rated their confidence in the choice, and received feedback. Although participants in both conditions showed equivalent levels of implicit learning, those under threat were delayed in becoming confident that learning had taken place. This inaccurate awareness of learning had consequences for post-task perceptions of performance and judgments of ability on future tasks. Discussion centers on the role of stereotype threat in hindering awareness of one’s abilities and the impact that might have on decision-making and motivation.
Preface

This research was conducted with the permission of the University of British Columbia Office of Research Studies Behavioural Research Ethics Board, certificate number H10-00896.
# Table of contents

Abstract................................................................................................................................. ii  
Preface ................................................................................................................................. iii  
Table of contents.................................................................................................................... iv  
Acknowledgements ............................................................................................................... vi  
Chapter 1: Introduction ......................................................................................................... 1  
1.1 Stereotype threat – overview ......................................................................................... 2  
1.2 Self-perception ................................................................................................................ 4  
1.3 Stereotype threat and self-perception ............................................................................. 11  
Chapter 2: Study .................................................................................................................... 18  
2.1 Method .......................................................................................................................... 18  
2.1.1 Participants and design .............................................................................................. 18  
2.1.2 Procedure ................................................................................................................ 18  
2.1.3 Measures .................................................................................................................. 20  
2.1.3.1 Pre-task anxiety .................................................................................................... 20  
2.1.3.3 Post task perception of learning ........................................................................... 21  
2.1.3.4 Post task motivation ............................................................................................. 22  
2.1.3.5 Stereotype related anxiety ..................................................................................... 22  
2.1.3.6 Manipulation check ............................................................................................... 22  
2.2 Results .......................................................................................................................... 23  
2.2.1 Analytic strategy ....................................................................................................... 23  
2.2.2 Manipulation check ................................................................................................... 23  
2.2.3 Anxiety .................................................................................................................... 23  
2.2.4 Objective categorization performance ...................................................................... 24  
2.2.5 Subjective perception of performance .................................................................... 25  
2.2.6 Relative accuracy of FOK rating .............................................................................. 27  
2.2.7 Post-task assessments of performance ..................................................................... 28  
2.2.8 Motivation .............................................................................................................. 31  
Chapter 3: Conclusion .......................................................................................................... 34  
3.1 Limitations and future directions .................................................................................. 38  
References ........................................................................................................................... 50  
Appendix: Sample Stimuli .................................................................................................... 61
List of Figures

Figure 1. Performance across the category learning task…………………………………45

Figure 2. Feeling of knowing ratings (adjusted for performance) across the categorization
task……………………………………………………………………………………………………46

Figure 3. Number of trials to reach the criterion for explicit learning, controlling for
implicit learning point……………………………………………………………………………47

Figure 4. Implicit learning point and explicit learning point predicting post task estimates
of learning…………………………………………………………………………………………48

Figure 5. Delayed confidence mediating stereotype threat induced under-confidence in
future performance…………………………………………………………………………………49
Acknowledgements

First and foremost, I would like to thank my advisor, Toni Schmader, for her selfless guidance and support over the past two years. Thank you also to my colleague, Alyssa Croft, for her mentorship, critical insights, and friendship. This thesis would have been impossible without both of you.
Chapter 1: Introduction

Accurate self-knowledge can inform a number of meaningful questions about yourself: Who am I? Who was I in the past? Who will I be in the future? What are my strengths and weaknesses? What career should I pursue? By the same token, these self-assessments have numerous significant consequences. Explicit self-views have been shown to predict behaviour on a number of important outcomes (Funder & Colvin, 1991). Accurate self-knowledge is used to guide the goals that a person pursues (Gollwitzer, 1999). Furthermore, misjudgments in self-knowledge have been shown to result in missed opportunities and the pursuit of inappropriate goals (Ehrlinger & Dunning, 2003). In fact, the ability to form accurate self-knowledge is thought to be so vital that researchers suggest it represents a predominant component of intelligence (Sternberg, 1996). So imagine the implications if the ability to accurately self-assess was systematically undermined for certain segments of the population.

There is now a growing body of evidence suggesting that stigmatized individuals might develop inaccurate self-assessments. For example, Aronson and Inzlicht (2004) demonstrated that African American students, particularly those most concerned about begin negatively stereotyped, were miscalibrated in their performance estimates and had unstable academic self-efficacy over the course of several days. Furthermore, a number of meta analyses have revealed that girls show comparable math performance to boys, but are significantly less confident in their math abilities, have lower self-efficacy, and have a weaker relationship between their math self-concept and achievement (Hyde, 2004; Chui & Klassen, 2010; Eccles, 1994; Pajares, 2005; Rittmayer & Beier, 2010). This inability to form accurate self-assessments has consequences for their perceptions of performance.
and their motivation to enter into certain domains (Ehrlinger & Dunning, 2003; Critcher & Dunning, 2009). Thus, over the past 10 years, evidence has continued to mount suggesting that stigmatized individuals can have difficulty assessing their true potential. However, the cognitive processes underlying these inaccurate self-perceptions have remained unexplored. The goal of the present research is to explore how facing a negative stereotype can lead to inaccuracies in self-perception.

In following sections I will provide a brief over of the empirical study of stereotype threat. I will then outline research on self-perception and how it can be rendered inaccurate. Following that, I will turn back to stereotype threat and posit how processes elicited by stereotype threat could result in inaccurate self-perception.

1.1 **Stereotype threat – overview**

Steele and Aronson (1995) proposed that for a stigmatized individual, the mere knowledge that a stereotype exists and might be applied to them could result in underperformance on tests of academic ability. They called this phenomena stereotype threat. In their seminal paper, Steele and Aronson (1995) demonstrated how situational factors trigger stereotype threat for minority students. When told that a difficult standardized test measured intellectual performance, African American participants performed much worse than when told the same test did not measure their intellectual ability. This finding provided a situational explanation for the achievement gap between White and African American students in contrast to prevailing genetic explanations. It in fact pointed to the power of the situation and the psychological processes it cues in bringing about group differences in performance.
This original finding has since been shown to affect a number of different groups and to span many different situations. To name a few, stereotype threat has been shown to undermine performance for African Americans on a verbal test (Steele & Aronson, 1995), the performance of low socioeconomic status children on an intelligence test (Croizet & Claire, 1998), White men’s performance on a math test in which their performance will be compared to Asian men (Aronson et al., 1999), and women’s performance on a math test that is thought to reveal gender differences in ability (Spencer, Steele, & Quinn, 1999). It has also been shown that a broad range of situational cues can elicit stereotype threat (Murphy & Jones-Taylor, 2012. Some of the most generative work in the field has outlined the underlying processes behind stereotype threat. Examining the cognitive operating principles of the phenomena has revealed that stereotype threat highjacks cognitive resources required for optimal performance in tests of complex cognitive skills (Schmader, Johns, and Forbes, 2008). Taking a look at the biological processes behind the phenomena has revealed that stereotype threat results in neurobiological and physiological changes that lead to poor cognitive performance (Mendes & Jamieson, 2012).

Stereotype threat has its roots in explaining performance differences between minority and majority group members. Much of the research to date has focused exclusively on performance (see Inzlicht & Schmader, 2012 for review). However, the fear of academic inferiority might introduce other burdens to stigmatized students that have as of yet have been unexplored. That is, stereotype threat could shape processes that are independent of performance, but have important consequences related to academics. In this thesis, I will suggest that stereotype threat can undermine self-perception
processes in a performance situation. This work will fill an important gap in the research on stereotype threat by examining how blind spots in self-knowledge can be the result of the fear of confirming a negative stereotype. In order to examine this question, this research will integrate methods for assessing implicit learning and meta-cognition with a theoretical framework for understanding stereotype threat. The aim of this research will be to show that stereotype threat can hinder a student’s ability to acquire accurate self-knowledge about their performance, even in situations in which performance is unaffected.

1.2 Self-perception

The ability to acquire accurate self-knowledge is not a straightforward one. In fact, it has been shown to be an intrinsically difficult task (Dunning, Heath, & Suls, 2004). Models of self-perception have been proposed in order to outline how this task can be made easier and more difficult (Dunning, 2005; Telock, 1984; Hofmann & Wilson, 2010; Gawronski & Bodenhausen, 2011). I will begin my discussion of self-perception by outlining the mechanisms thought to underlie self-perception processes.

When making self-assessments people use information that is available to them to inform their inferences (Dunning, 2005; Wilson, 2002). Attribution theories deal with how people use information in their environment to draw conclusions about themselves and others (Fiske & Taylor, 1984). One of the most important contributions to attribution theory was Daryl Bem’s Self-Perception theory (1967, 1972). Self-perception theory looks to explain how people draw inferences about themselves. Bem (1967) argued that internal cues are often ambiguous or inaccessible, making them inadequate sources of information for self-assessment. According to self-perception theory, people infer their reactions, attitudes, and emotions by observing external
cues. That is, people infer their own attitudes by attending to their behaviors. Thus, external behaviors are posited to offer a rich source of information that people can use to make judgments about themselves. However, psychologists quickly discovered that people have very little insight into the meaning of their own behaviors (Nisbett & Wilson, 1977b).

The discovery that processes outside of conscious awareness guide much of our behavior raised the question of how well we can really know ourselves (Nisbett & Wilson, 1977; Wilson, 2002). Freud (1960) was the first psychologist to suggest that a vast amount of our cognition occurs outside of our awareness. Psychoanalytic theories were largely exiled from mainstream psychology for a number of years but experienced a rebirth with the introduction of dual process theories in the past three decades. Since then, research on dual processes theories has proliferated in psychology (Schneider & Shiffrin, 1977; Chaiken, 1980; Greenwald, 1992; Liberman, 2002; Strack & Deutsch, 2004). These models of information processing propose two processing systems (under many different names; see Gawronski and Creighton, 2012 for review) with different operating principles: (1) a controlled system and, (2) an automatic system. The controlled system is said to carry mental operations that are (a) initiated intentionally; (b) require cognitive resources; (c) can be stopped voluntarily and; (d) operate within conscious awareness (Bargh, 1994). The automatic system is thought to operate under principles that directly complement those of the controlled system. Automatic processes are (a) initiated unintentionally; (b) require few cognitive resources to operate; (c) cannot be stopped voluntarily and (d) operates outside of conscious awareness. For the purpose of our discussion, the key distinction is that automatic processes are not available to conscious awareness. That is, we cannot directly observe the operation of automatic processes.

However, more recently researchers have suggested that although we cannot directly
access the **operation** of automatic processes, the **output** of automatic processes can become available to consciousness (Deutsch & Strack, 2004; Wilson, 2002; Wilson & Dunn, 2004; Hofmann & Wilson, 2010; Ward; 2012). The processing outputs can take the form of conceptual activation, affective feelings, and incipient behavioral responses (Deutsch & Strack, 2004). These outputs are experienced by consciousness as having a distinct phenomenal quality. However, in order to be experienced, they must be attended to. Research has shown that attending to these phenomenal cues results in more accurate self-assessment of nonconscious dispositions. Gawronski and Lebel (2008) demonstrated this idea by manipulating participants’ focus of attention and then examining the relationship between implicit and explicit attitude measures. Participants were either instructed to focus on their affective reactions to an attitude object, or the reasons for liking that object. Participants who focused on the reasons that they liked an object showed no correspondence between their implicit and explicit attitudes. However, participants who were instructed to focus on their affective reactions to the object showed a substantial explicit-implicit correspondence. Furthermore, a recent meta-analysis found that the implicit-explicit correspondence is higher when the focus of attention is affective, rather than cognitive (Hofmann et al., 2005). Thus, by focusing attention towards feelings, a person can gain some insight into the nature of attitudes that are represented non-consciously. However, can people do the same to accurately assess other forms of self-knowledge? For instance, can we inform perceptions of learning or knowing in the same way?

Like attitudes, learning can happen non-consciously while our self-perception of learning reflects more conscious or explicit awareness of what we know. Much of our real world learning is thought to take place outside of awareness and independent of attention (Reber, 1993). This type of learning is called implicit learning (Reber, 1993). In implicit learning, information is
learned as associations between concepts (Strack & Deutsch, 2004; Gawronski & Bodenhausen, 2006). Implicit learning occurs slowly with repeated experience (McClelland, McNaughton, & O’Reilly, 1995). Unlike explicit learning, implicit learning happens independent of attention or motivation (McClelland, McNaughton, & O’Reilly, 1995), is not dependent on executive resources (e.g., Foerde, Poldrack, & Knowlton, 2007; Waldron & Ashby, 2001), is unaffected by consciously held learning goals (DeShon & Alexander, 1996), and is not impaired by general anxiety or stress (e.g., Kirschbaum, Wolf, Wippich, & Hellhammer, 1996; Lupien et al., 1997). Even though implicit learning happens outside of awareness, it elicits information that if attended to can inform self-assessments of learning (Topolinski & Strack, 2009; Deutsch & Strack, 2004).

Research examining self-assessments of ability and learning has been broadly studied by cognitive psychologists under the topic of meta-memory (Nelson & Narens, 1990; Koriat, 1993; Metcalfe, 1998). Meta-memory is defined as the knowledge, monitoring, and control of learning and memory processes. Monitoring of knowing has typically been measured by asking people to make feeling of knowing ratings (FOK; Hart, 1965, Koriat, 1993). FOK ratings assess subjective feelings regarding whether or not knowledge exists within short-term or long-term memory. FOK ratings are used to measure participants’ future ability to remember an item that is currently not available to conscious recall, but has previously been learned (Metcalfe, 1993). FOK ratings have been shown to be derived from partial activation of the to-be recalled information (Koriat, 1993; Schwartz, 2002). This activation results in processing fluency that is then used to inform the FOK judgment (Koriat, 1993; Reber & Schwarz, 1999). Previous research has shown that participants’ FOK judgments are relatively accurate in their predictions of future recall (Metcalfe, 1993; Schacter, 1983). Furthermore, there is a large body of research showing a negative correlation between FOK ratings and the amount of time spent studying (see Son & Metcalfe,
Thus, these meta-cognitive judgments are not just fleeting feelings, but have consequences for real behaviour (Metcalfe, 2009). Yet, how do these feelings come into conscious awareness?

Over the course of an implicit learning task, a chain of events is assumed to take place that generate an output that can be used to inform FOK judgments about learning (Topolinski & Strack, 2009; Koriat, 1993; Reber & Schwarz, 1999). First, associations are established between perceptual information and the correct responses (Ashby & O’Brien, 2004). Associations between stimulus dimensions and the correct responses are established by attending to feedback in the environment. As learning takes places, the strength of the associations between perceptual features and the correct response are strengthened (Ashby & O’Brien, 2004, Topolinski & Strack, 2009). Furthermore, the correct responses can become linked with behavioral schema that elicit behavioral outputs (Deutsch & Strack, 2004). Finally, as learning take place, activation begins to spread between perceptual features, the correct response, and behavioral schema with increasing fluency (Topolinski & Strack, 2009). Three types of internal cues are thought to be elicited by implicit learning: (1) conceptual activation, (2) feelings of fluency and, (3) behavioral responses. These internal cues have distinct phenomenal qualities and, when attended to, can inform metacognitions about learning (Koriat, 1993; Reber & Schwarz, 1999; Topolinski & Strack, 2009). However, the ability to attend to these phenomenal cues is gated by attentional resources (Hoffman & Wilson, 2010; Souchay, Isingrini, Clarys, Taconnat, & Eustache, 2004; Widner, Otani, & Winkelman, 2005).

In sum, self-perception takes place when we carefully observe our immediate experience. This allows us to gain some insight into the vast amount of cognition that is taking place outside of our awareness. When attention is directed in the right place we
can come to know ourselves and our abilities a little bit better. However, a wealth of evidence suggests that these processes can be derailed by a number of situational and intrapsychic factors (Dunning, 2005; Wilson, 2002; Nisbett & Wilson, 1977). Two different ways have been proposed by which self-perception can become inaccurate: informational deficits and motivated cognition. I will discuss each of these in the sections that follow.

The accuracy of self-assessment is dependent upon the quality and quantity of the information available to someone (Dunning, 2005; Hofmann & Wilson, 2010). Thus, blind spots in self-knowledge can be the result of a simple lack of information. Personality psychologists have examined which traits people are able to accurately self-assess. The information available to someone about a given traits has been shown to moderate the ability to accurately assess that trait (Vazire, 2010). For example, people are more accurate at assessing their own internal traits (i.e. neuroticism); whereas, outside observers are more accurate at assessing external traits (i.e. extraversion; Vazire, 2010). This is thought to be the result of the advantaged informational vantage point we possess when assessing our internal traits (Vazire, 2010). That is, for internal traits we can easily attend to the distinct phenomenal and cognitive qualities of the trait; for external traits, outside observers have better sources of information from which to draw inferences. This runs somewhat contrary to self-perception theory. However, Bem (1972) argued that self-perception processes are most likely to rely upon external behavior when internal cues are weak or ambiguous and when external sources of feedback about attitudes are absent. This may be correct, however, attending to your external behaviors as a source of self-knowledge offers you less insight than turning to a friend and asking for their opinion!

Some people have inaccurate self-assessments because they lack competence in that
domain (Kruger & Dunning, 1999). Somewhat ironically, being incompetent in a domain makes it very difficult to accurately self-assess ability. This is because the ability to self-evaluate is dependent on the same knowledge or skills required for competent performance. For example, the ability to evaluate the grammatical correctness of this sentence is dependent upon knowing correct grammar. As an empirical demonstration of this phenomena, researchers have shown that college students who score in the bottom 25% in an exam think that they have outperformed the majority of the other test takers (Dunning, Johnson, Ehrlinger & Kruger, 2003). These blindspots can be overcome by providing students with accurate information to base their self-assessments. For instance, giving students accurate feedback allows them to accurately decide what they do and don’t know. Dunning and Captuto (2005) showed that graduate students’ self-assessments of their research methods knowledge was not correlated with their performance on a subsequent test. However, when students were provided feedback about the errors of omission they made, their self-assessments then correlated with their objective performance.

Thus, maybe not surprisingly, the accuracy of self-assessment is dependent upon the data people have available to them when making these judgments (Dunning, 2005; Hofmann & Wilson, 2010). If a person is not provided the necessary information their self-assessments will suffer.

Motivated cognitive processes are another source of error in self-assessment (Dunning, 1999). These processes can result in blind spots in self-knowledge that lead to flawed self-assessment. The majority of research examining these biases looks at how people’s desire to maintain their self-worth affects self-perception (Greenwald, 1980; Paulhus & John, 1998; Dunning, 2005). For example, research has shown that when your sense of self-worth is linked to a trait, you have very little insight into that trait.
Intelligence; Vazire, 2010). In fact, your friends have greater insight into your evaluative personality traits (Vazire, 2010; Epley & Dunning, 2004). There has been extensive research on how people distort information in order to maintain a positive self-view (see Dunning, 2005 for review). However, there is also growing evidence that people will also work to preserve overly negative self-views (Swann, 1997). This has led some researchers to suggest that information is often disregarded when it does cohere with other self-beliefs (Festinger, 1957; Heider, 1958). This can lead to useful information being ignored, resulting in blindspots in self-knowledge. For example, positive feedback given during a math class might be disregarded if a person has a chronically low math self-concept.

In sum, people often make inaccurate self-assessments because they fail to detect and utilize of relevant information. (Dunning, 2005; Wilson & Nisbet, 1997; Wilson, 2002). Deficits in self-perception can be introduced by a disadvantaged informational vantage point (Vazire, 2010), incompetence (Dunning et al., 2003), a lack of useful feedback (Dunning & Captuto (2005), attentional biases (Gawronski and Lebel, 2008), and motivated cognition (Helzer & Dunning, 2012). I will now turn my attention back to stereotype threat and outline how the psychological processes it elicits could derail self-perception.

1.3 Stereotype threat and self-perception

Steele and Aronson’s (1995) demonstration that situational cues can undermine performance for minority students prompted a wave of research showing that stereotype threat could be experienced by a number of different groups on a number of different performance tasks. The mechanism behind stereotype threat was, initially, largely
unexplored. In an effort to outline the mechanisms thought to underlie the performance
detriments resulting from stereotype threat, Schmader, Johns, and Forbes (2008) laid out
the integrated process model.

The integrated process model identified working memory as a central mediating
mechanism behind performance decrements experienced due to stereotype threat. A
number of pathways are theorized to reduce working memory. These include a
physiological stress response, a meta-cognitive monitoring process, and active regulation
of negative thoughts. Theorizing from this model, I am going to suggest that stereotype
threat produces informational deficits and/or motivational biases that can hinder self-
perception. I will consider both of these deficits through lens of what we know about
stereotype threat in order to outline how stereotype threat can undermine self-perception.

As I have already reviewed, informational deficits can be the result of an inability
to attend to useful information. Attentional focus is the responsibility of working memory.
Working memory is a limited capacity system that temporarily stores information
(Baddeley, 2003). Working memory acts to store and combine information, facilitating a
range of cognitive abilities, such as reasoning, learning and comprehension (Baddeley,
2003). Schmader and Johns (2003) provided the first empirical demonstration that
minority students and women exhibit lower working memory capacity when experiencing
stereotype threat. Follow-up research has demonstrated that stereotype threat undermines
working memory by interfering with attentional focus (Hutchison, Smith, & Ferris, 2012;
Mrazek et al., 2011).

One of the mechanisms by which stereotype threat undermines attentional focus is
through the suppression of unwanted negative thoughts and feelings. Stereotype threat
elicits feelings of anxiety that might be perceived as signaling poor task performance (Schmader et al., 2009). Thus, individuals under stereotype threat are hypothesized to engage in effortful suppression processes in order to banish unwanted cognitions. Johns, Schmader, and Inzlicht (2008) demonstrated reduced working memory in a stereotype threat context was mediated by avoidance of anxiety related cues. Furthermore, research employing fMRI has also shown that stereotype threat results in activation in regions of the brain implicated in emotion regulation (Krendl, Richeson, Kelley, & Heatherton, 2008; Wraga et al., 2007).

In a performance situation, stigmatized students’ attentional resources are being devoted to something other than the task at hand (Mrazek et al., 2011; Beilock, Rydell, & McConnell, 2007; Johns, Schmader, & Inzlicht, 2008). Research has shown that impaired executive function leads to inaccurate perceptions of performance (Souchay et al., 2004) and a reduced ability to attend to immediate conscious experience (Sayette, Reichle, & Schooler, 2009; Sayette, Schooler, & Reichle, 2010) Thus, individuals who are facing stereotype threat have a naturally occurring cognitive load that draws their attention away from their performance on the task. This attentional bias might result in individuals being unable to attend to informational cues (i.e. conceptual activation, feelings of fluency, and behavioral responses) that would otherwise inform their perceptions of performance.

The second way that self-perception can be undermined is through motivational biases. Stereotype threat elicits a number of motivational constraints that could lead to erroneous self-assessments. The integrated process model suggests that stereotype threat elicits three cognitions: (1) I do well in domain X, (2) I am a member of group G (3) group G is not expected to do well in X (Schmader et al., 2008). This cognitive
imbalance results in feelings of self-doubt and uncertainty. A number of studies have documented increased feelings of uncertainty during stereotype threat (Beilock, Rydell, & McConnell, 2007; Steele & Aronson, 1995; Cadunu, Mass, Rosabianca, & Kiesner, 2005). For instance, Cadunu and colleagues (2005) showed that women who completed a math test under stereotype threat reported having more negative math-related thoughts than women in the control group. Furthermore, Beilock et al., (2005) showed that women who were reminded of a negative stereotype about their performance in math showed reduced performance only on math problems that relied heavily on verbal working memory resources. The authors suggested that ruminations related to self-doubt could be taxing the phonological loop, which lead to the reduced performance on verbal math problems.

The feelings of uncertainty elicited by stereotype threat are hypothesized to drive a meta-cognitive monitoring response that leads individuals to become vigilant for cues that provide information about their performance. Because this vigilance is colored by a desire to disconfirm the stereotype, this results in greater attentiveness to signs of failure. Schmader and colleagues’ (2008) theorizing is supported by a number of empirical findings (Johns, Schmader, & Inzlicht, 2008; Kaiser, Vick, & Major, 2006; Murphy Steele, & Gross, 2007; Seibt & Forster, 2004). Seibt and Forster (2004) demonstrated that stereotype threat alters an individual’s regulatory focus such that they become overly vigilant towards the possibility of negative performance. Furthermore, a recent study demonstrated that minority students, in a stereotype threat context, are automatically more vigilant to errors on a performance task (Forbes, Schmader, & Allen, 2008).
Thus, self-doubt and vigilance towards negative performance cues could change a person’s motivational orientation. This could result in over-attending to negative information, making it difficult to process performance cues that connote competency. Research has shown that activating motivational orientations can lower the threshold for processing information that coheres with the activated orientation (Strack & Deutsch, 2004). For example, Neumann and Strack (2000) demonstrated that positive words were processed more readily when a person adopted an approach orientation. On the other hand, negative words were processed more easily when participants adopted an avoidance motivation. Thus, by activating an avoidance motivation (Seibt & Forster, 2004) and increasing vigilance towards negative performance feedback (Forbes et al., 2008), stereotype threat could raise the processing threshold for informational cues that suggest competent performance. This could make it difficult for stereotyped individuals to process information that suggests they are performing well.

In sum, stereotype threat introduces a number of attentional and motivational biases that could undermine self-perception. But where might these blindspots in self-knowledge be most detrimental in an academic setting? Much of stereotype threat research has concentrated on examining how test performance can be undermined by stereotype threat. Performance in a testing situation obviously represents an important aspect of academics as it can dictate much of a student’s success (i.e. standardized testing, entrance exams, etc.). However, a much larger amount of time in the classroom is spent learning new information. That is, students spend a great deal of time developing knowledge and skills that they must later call upon in a testing situation. Furthermore, accurate self-assessments of learning are crucial because students regulate their study
habits based on these judgments (Son & Metcalfe, 2000; Metcalfe, 2009). Thus, the focus on our research was to examine how situations of stereotype threat impair self-assessments of learning.

Only a handful of studies have sought to examine stereotype threat effects on learning. Some of these studies demonstrate that stereotype threat can impair the ability to learn new information (Rydell, Shiffrin, Boucher, Van Loo, & Rydell; 2010; Rydell, Rydell, & Boucher, 2010; Jones-Taylor & Walton; 2010). However, this might not always be the case. Stereotype threat hinders performance by undermining working memory (Schmader & Johns, 2003). Therefore, learning that is not dependent upon working memory should not be affected by stereotype threat. As mentioned earlier, implicit learning has been shown to occur independently of executive function (Waldron & Ashby, 2001). Thus, implicit learning should not be impeded by stereotype threat. This presents an opportunity to examine the effect of stereotype on self-perception in a context were performance should be unaffected. In doing so, I am creating a situation that is analogous to the meta-analyses reported earlier in which performance is equivalent between minority and non-minority students, but self-perceptions of ability are more negative for minority students.

Thus, in present research I will employ an implicit learning task to examine how stereotype threat affects assessments of learning in a situation where learning is not undermined. Through the combination of an implicit learning task and meta-cognitive judgments of learning I can look for disassociations between actual learning and perceptions of learning. More specifically, by employing an implicit learning task I can hypothesize that learning should be unaffected by stereotype threat. On the other hand,
the ability to monitor performance should be undermined. These predictions can be made given the attentional requirements of each task. Implicit learning has been shown to be unaffected by cognitive impairments (Waldron & Ashby, 2001). The ability to monitor performance is however governed by attentional resources (Hoffman & Wilson, 2010; Souchay, Isingrini, Clarys, Taconnat, & Eustache, 2004; Widner, Otani, & Winkelman, 2005). To test these predictions, female participants completed an implicit category learning task under stereotype threat or control conditions while making online meta-cognitive judgments of learning. I hypothesize that participants in the stereotype threat condition will have difficulty monitoring their learning, which will result in inaccurate representations of learning during the task. These inaccurate representations will then have consequences for their post task perceptions of their performance and motivation.
Chapter 2: Study

2.1 Method

2.1.1 Participants and design

Sixty-seven female participants were recruited from the University of British Columbia’s Human Subject Pool. Because a test of our hypotheses requires participants to demonstrate successful learning of a categorization rule at some point during the task, participants who did not achieve the criterion for implicit learning (operationalized as eight consecutive correct responses) were excluded from further analyses (n = 11, see Waldron & Ashby, 2001, for precedent). There was no difference by condition in the number of people who failed to reach criterion, $\chi^2 = .875$, $p=.350$. Furthermore, two participants were excluding for explicitly mentioning stereotype threat during the funnel debriefing, and two participants were excluded for having average responses latencies over two standard deviations above the mean. This resulted in a final sample of 52 participants. Participants were randomly assigned to either a stereotype threat or control condition.

2.1.2 Procedure

Participants were recruited to take part in a study about learning. Participants were run in groups of 1 to 4. Participants were seated in individual cubicles. All instructions were delivered by a male experimenter. Participants first received instructions describing the category learning task. The sequence and purpose of the category learning task was described in detail.
Following the description of the category learning task, participants were then randomly assigned to condition. In the stereotype threat condition, the experimenter described the category learning task as measuring gender differences in spatial ability. To link their gender identity to the task, participants in the stereotype threat condition were then asked to indicate their gender prior to beginning the task. In the control condition, participants were told that the researchers were interested in examining how different features of the stimuli affected implicit learning. In the control condition, participants were not asked to indicate their gender prior to beginning the task. Thus, similar to previous research (Steele & Aronson, 1995), the manipulation of threat either framed the task as diagnostic of a stereotype relevant ability or as concerning features of the task itself.

After the stereotype threat manipulation, participants completed a self-report measure of anxiety before beginning the categorization task. On each trial of the categorization task (described in more detail below), participants made a category choice, rated their confidence in their choice (a common measure of feeling of knowing, Koriat, 1993), and received feedback on their choice. Following the categorization task, participants completed a series of self-report measures designed to assess their future motivations, perceptions of performance, and basic demographic information before being fully debriefed. During the debrief participants were informed about the deception used in the stereotype threat condition.
2.1.3 Measures

2.1.3.1 Pre-task anxiety

To provide a composite measure of pre-task anxiety (α = .84), participants rated the extent to which they felt agitated, anxious, nervous, uneasy, and worried before completing the categorization task (1 = not at all, 7 = extremely).

2.1.3.2 Category learning task

The measure of implicit learning was adapted from one used by Waldron and Ashby (2001). In the standard category learning task, participants classify stimuli into category “A” or category “B” by the appropriate key on the keyboard. They start the task with no knowledge of the categorization rule but gradually improve through trial and error learning. Following each categorization, participants receive feedback about their category choice. In the present study we modified the task such that following each categorization, participants made a feeling of knowing rating (FOK). Specifically, they were asked the following “How confident are you in your category choice?” and selected a response from 1 (Very unconfident) - 7 (Very confident). After the feeling of knowing rating, participants received feedback about their category choice. Participants completed 192 trials that included a category choice, feeling of knowing rating, and feedback.

The categorization stimuli were taken from Waldron and Ashby (2001) and consisted of colored shapes presented on a black background (see Appendix for stimuli). The stimuli varied on four dimensions: shape (square or circle), color (blue or yellow), number (one or two), and size (big or small). Sixteen unique stimuli were constructed by the factorial combination of the four binary dimensions: color, size, shape, and number. Three implicit categorization rules were constructed involving two of the four...
dimensions, and participants were randomly assigned to one of the three categorization rules. Categorization rules were constructed in the following manner. First, one level of each stimulus dimension was arbitrarily assigned the numerical value of -1, and the other level was assigned a value of +1 (e.g., blue = -1 and yellow = +1). The stimuli were then assigned to categories according to the following rule:

If \( \text{value}(X) + \text{value}(Y) \neq 0 \)

assign to category A

else, assign to category B.

This resulted in three categorization rules (e.g., if the stimulus is blue and big or yellow and small then the item belongs to category A; if the stimulus is blue and small or yellow and big then the item belongs to category B). The categorization rules were constructed so that participants would only achieve chance performance if they applied sub-optimal categorization rules (i.e., simple 1 or 2 dimensional categorization rules).

2.1.3.3 Post task perception of learning

Participants completed four items designed to assess post task perception of learning. Participants indicated the point during the categorization task at which they thought they knew the category rule. This was reported using a slider scale (0: beginning of the task – 100: end of the task). Participants were then asked to explicitly report the categorization rule. Participants rated how confident they were that the categorization rule they reported was correct (1: Very unconfident - 7: Very confident). Finally, participants estimated
how many errors (percentage) they made during the task.

2.1.3.4 Post task motivation

To assess post task motivation, participants were asked to help pilot test a new categorization task by completing a number of trials of their choosing. Participants were provided with a description of the task and asked to indicate how many trials they would be willing to complete (0 -200) and how long they thought it would take them to learn the categorization rule. Participants never completed the pilot test.

2.1.3.5 Stereotype related anxiety

Participants rated three items designed to assess anxiety related to confirming a negative stereotype. On a 7-point scale, participants rated their agreement with the following items (α = .77), “I worry about being seen as stereotypical of my gender in terms of how I did on this category learning task.”; “I worry that if I performed poorly on the category learning task, the experimenter will think that it's because of my gender”, and “I worry that I might have done poorly on the category learning task because I was anxious about confirming the stereotype about my gender on this task.” Responses to these three items were averaged to create an index of anxiety related to stereotype threat.

2.1.3.6 Manipulation check

As a manipulation check, participants rated their agreement with three statements (α = .73), “I think the experimenter probably believes that gender is predictive of performance on the computer task I completed in today’s study ”, “Regardless of what you personally believe, there is a general stereotype about gender and spatial ability:”, and “I believe that
gender is predictive of spatial ability.” Participants responded using a 7 point scale (1 – women will out perform men; 7 – men will outperform women). Responses to these three items were averaged to create an index of stereotype threat based concern.

2.2 Results

2.2.1 Analytic strategy
Unless otherwise specified, all analyses were conducted with independent samples t-tests (condition: stereotype threat vs. control).

2.2.2 Manipulation check
To establish that the manipulation cued stereotype threat, I first analyzed effects on the manipulation check items. As predicted, participants in the stereotype threat condition showed greater concern about gender stereotypes ($M_{\text{threat}} = 4.47$, $SD = .96$), than did participants in the control condition ($M_{\text{control}} = 3.94$, $SD = .97$), $t(50) = 2.01$, $p = .049$.

2.2.3 Anxiety
To assess whether participants in the stereotype threat condition experienced greater pre- or post-task anxiety, I examined the self-reported anxiety measures. For pre-task anxiety, the analysis yielded no main effect of condition ($M_{\text{threat}} = 2.54$, $SD = .78$; $M_{\text{control}} = 2.70$, $SD = .84$), $t(50) = -0.72$, $p = .48$. The same pattern was found for post-task anxiety ($M_{\text{threat}} = 2.44$, $SD = 1.11$, $M_{\text{control}} = 2.09$, $SD = .94$), $t(50) = 1.21$, $p = .231$. This pattern of results is not uncommon in stereotype threat research. Past research has shown that stereotype threat seldom leads to conscious reports of elevated anxiety and actually elicits suppression of anxiety-related thoughts and feelings (Johns et al., 2008).
2.2.4 Objective categorization performance

Although we did not anticipate that stereotype threat would impair learning given the implicit nature of learning on this task, I first analyzed whether stereotype threat affected performance on the categorization task. I conducted a number of analyses looking at different indices of performance. First, I examined whether participants’ learning across the course of the categorization task differed by condition. The categorization task was divided into 16 blocks of 12 trials. The percentage of correct categorizations was calculated for each block. A 2 (condition: stereotype threat vs. control) x 16 (block) mixed model ANOVA, with repeated measures on the last factor, was conducted to examine whether participants’ performance over the course of the task differed by condition. The analysis revealed that participants’ performance did not differ by condition, $F(1, 50) = .37, p = .532$. Furthermore, across conditions, participants’ performance across blocks was fit by a linear function, $F(1, 51) = 166.74, p < .001$. The block x condition interaction was not significant, $F(1, 51) = 1.28, p = .263$. These results are consistent with the notion that participants in both conditions were gradually improving their performance as the task progressed (see Figure 1).

Next, I assessed how long it took participants to reach the criterion (8 correct categorizations in a row) for implicit learning. On average participants reached the criterion within 75 (out of 192) trials. Implicit learning criterion scores ranged from 11 to 188. Participants’ performance was not affected by the stereotype threat manipulation ($M_{threat} = 73.42, SD = 47.82; M_{control} = 75.38, SD = 52.05$), $t(1, 50) = -0.14, p = .89$.

I also examined whether overall performance (percent correct) on the categorization task differed by condition. The analysis revealed that stereotype threat did
not undermine performance on the categorization task ($M_{\text{threat}} = .77, SD = .11; M_{\text{control}} = .79, SD = .13$), $t(50) = -0.64, p = .52$.

Finally, I assessed whether in the stereotype threat condition, fewer participants learned the rule by the end of the task. To test this, I looked at participants’ descriptions of the categorization rule at the end of the task. In the stereotype threat condition, 77% of participants accurately described the categorization rule; in the control condition, 73% of participants did the same. There was no difference by condition in the number of people who failed to accurately describe the categorization rule, $\chi^2 = 1.0, p = .75$. These analyses all demonstrate that, consistent with our expectation, stereotype threat did not undermine performance on the implicit category learning task.

### 2.2.5 Subjective perception of performance

I conducted a similar set of analyses to test our primary hypothesis that participants’ subjective perception of performance would be affected by stereotype threat. Unless otherwise specified, all analyses were conducted controlling for the parallel measure of objective performance. In other words, for each analysis I covaried the index of objective performance that was most consistent with the dependent variable (e.g., when examining overall confidence I controlled for overall performance). This allowed me to test participants’ perceptions of performance, controlling for individual variation in actual performance.

First, I conducted a series of tests to assess whether participants experienced lower levels of confidence in the stereotype threat condition. In the first analysis, I examined the degree to which participants’ FOK ratings under or overestimated their actual performance across the course of the categorization task. To assess this, I divided...
the task into 16 blocks of 12 trials. For each block, I subtracted the percentage of correct responses from the proportion of FOK scores above the midpoint. Thus, positive values indicated overconfidence, negative values indicated underconfidence, and zero indicates accuracy. A 2 (condition: stereotype threat vs. control) x 16 (block) mixed model ANOVA, with repeated measures on the last factor, was conducted on participants under/overconfidence scores. This analysis revealed that across conditions, participants’ performance across blocks was fit by a linear function, $F(1,50) = 30.71, p = .001$. However, under/overconfidence scores did not differ by condition, $F(1,50) = .07, p = .793$. The block x condition interaction was also not significant, $F(1,50) = 0.31, p = .579$. These results are consistent with the notion that participants in both conditions were gradually developing confidence in their performance as the task progressed (see Figure 2). Although this analysis does not show evidence that stereotype threat affected FOK ratings, I turned to what might be a more sensitive test of our hypothesis.

I next tested whether participants in the stereotype threat condition took longer to develop awareness of their objective learning. To test this, I examined the point at which participants reached a criterion for explicit awareness of learning (eight consecutive confidence ratings above the scale midpoint), while controlling for the point at which they reached the criterion for implicit learning (which was a significant covariate, $F(1,49) = 16.41, p < .001$). This analysis allowed me to examine whether stereotype threat hindered participants’ ability to become consciously aware of their objective performance on the task. The analysis revealed a main effect of condition, $F(1,49) = 4.83, p = .03$, such that participants in the threat condition took more trials to become consciously aware of what they had learned implicitly ($M_{threat,adj} = 114.38, SE = 7.93$; $M_{control,adj} = 89.74, SE =$
7.93). This result is consistent with our hypothesis that stereotype threat makes it difficult to consciously attend to performance (see Figure 3).

Finally, I examined whether overall confidence differed by condition. Participants’ FOK scores were averaged across the categorization trials to provide an index of overall task confidence. To examine whether participants in the threat condition were less confident in their task performance, I conducted a one way (condition: stereotype threat vs. control) ANCOVA, controlling for overall performance (which was a significant covariate, $F(1,49) = 49.40, p < .001$). The analysis revealed no difference in overall task confidence ($M_{\text{threat}} = 5.51, SE = .12 \ M_{\text{control}} = 5.62, SE = .12$), $F(1,49) = .469, p = .497$. Thus, it appeared that participants were not experiencing lower confidence in the stereotype threat condition. These analyses suggest that delayed perceptions of performance in the stereotype threat condition are not the result of overall lower confidence in performance across the entire task.

2.2.6 Relative accuracy of FOK rating

In the metacognition literature, gamma correlations have been used to measure the accuracy of FOK judgments (see Nelson, 1994). Thus, gamma correlations were calculated to assess the relationship between participants’ FOK ratings and accuracy of their categorizations. A correlation was calculated for each participant by averaging across the 192 trials. Higher correlations represent a higher degree of accuracy between meta-cognitive judgments and performance on the categorization task. Contrary to predictions, there was no difference in the accuracy of participants’ metacognitive judgments, ($M_{\text{threat}} = .42, SD = .27; M_{\text{control}} = .44, SD = .29$) $t(50) = -.321, p = .750$. In both conditions, the gamma scores were significantly different from zero, $t(51) = 10.97, p$
< .001. This suggests that, in both conditions, participants could accurately assess whether they knew the categorization rule.

I also examined participants’ gamma correlations before they reached the criterion for explicit learning. This allowed me to look at the correspondence between actual performance and perception of performance before participants became consistently confident of the categorization rule. This analysis also revealed no significant condition differences in the accuracy of participants’ metacognitive judgments, \( M_{\text{threat}} = .26, SD = .26; M_{\text{control}} = .26, SD = .31 \) \( t(45) = -0.04, p = .97 \). Furthermore, in both conditions, the gamma scores were significantly different from zero, \( t(46) = 6.37, p < .001 \).

2.2.7 Post-task assessments of performance

Next, I tested how stereotype threat affected post-task assessments of performance. All analyses were conducted controlling for objective performance. The index of objective performance again changes depending upon the analysis. For each analysis I chose the index of objective performance that was most consistent with the dependent variable of index (e.g., when examining estimated errors I controlled for overall performance). This allowed me to examine participants’ perceptions of performance, independent of actual performance.

I first conducted a series of tests to examine whether stereotype threat led to more negative perceptions in performance as a whole. I examined whether there were condition differences in the number of errors participants estimated they made on the categorization task. I conducted a one-way (condition: stereotype threat vs. control) ANCOVA, controlling for overall performance (which was a significant covariate, \( F(1,49) = 76.05, p < .001 \)), to assess whether stereotype threat led participants to overestimate the number of
errors that they made during the categorization task. The analysis revealed no significant
difference in post-task estimations of errors ($M_{\text{threat}} = 35.42$, $SD = 3.84$; $M_{\text{control}} = 36.77$,
$SE = 2.84$), $F(1,49) = 0.11$, $p = .737$.

I then examined whether there were condition differences in participants’
confidence in the rule they reported at the end of the categorization task. Post-task
certainty scores were submitted to a one-way (condition: stereotype threat vs. control)
ANCOVA, controlling for implicit learning point (which was a significant covariate,
$F(1,49) = 7.45$, $p = .009$). The analysis revealed no significant difference in post-task
confidence ($M_{\text{threat}} = 5.83$, $SE = .38$; $M_{\text{control}} = 5.44$, $SE = .38$), $F(1,49) = .53$, $p = .472$.

Finally, a one-way (condition: stereotype threat vs. control) ANCOVA,
controlling for implicit learning point (which was a significant covariate, $F(1,49) = 39.20$,
$p < .001$), was conducted to assess whether participants in the stereotype threat condition
estimated learning the categorization rule later than participants in the control condition.
The analysis revealed no significant difference in learning point estimations ($M_{\text{threat}} =
45.61$, $M_{\text{control}} = 38.88$), $F(1,49) = 1.88$, $p = .17$. From these analyses, it appears that
stereotype threat did not result in more negative post-task perceptions of performance.

Although there were no mean differences in performance perceptions, there might
be differences in the process by which these judgments are made. Next, I examined what
information participants were using to inform their beliefs about their performance.
Specifically, I examined whether implicit and explicit learning points made unique
contributions to participants’ post-task perceptions of when they learned the rule. If
participants have access to the point at which they developed implicit learning, then that
information should predict their post-task estimates of learning. However, if participants
have difficulty gaining insight into implicit learning, their subjective FOK ratings might be more predictive of their post-task assessments. Thus, I hypothesized that among participants in the control condition, the point at which they reached criterion for implicit learning would predict their self-perceptions after the task; whereas, among participants in the stereotype threat condition, the point at which they reached the criterion for explicit awareness of learning would predict their post-task self-perceptions.

To test this hypothesis, I conducted a hierarchical regression in which implicit learning point, explicit learning point, and condition were entered into the first step, the two-way interaction terms were entered into the second step, and the three-way interaction was entered into the third step predicting participant’s post-task estimated learning point. The analyses revealed a significant main effect of implicit learning point, $\beta = .38, t(45) = 2.70, p = .01$, and explicit learning point, $\beta = .40, t(45) = 2.76, p = .008$, and a significant interaction between implicit learning point and explicit learning point, $\beta = -.29, t(45) = -2.85, p = .007$. These effects were all qualified by a significant three-way interaction between implicit learning point, explicit learning point, and condition, $\Delta R^2 = .06, \beta = .39, t(45) = 3.09, p = .003$.

Separate multiple regression analyses were then conducted for each condition in which participants’ post task learning point estimate was regressed onto their implicit and explicit learning points simultaneously. As hypothesized, in the control condition, implicit knowledge of the categorization rule predicted post-task estimates of learning, $\beta = .79, t(23) = 3.24, p = .004$; explicit awareness was not a significant predictor, $\beta = -.11, p > .65$. In the threat condition, explicit knowledge of the categorization rule predicted post-task estimates of learning, $\beta = .73, t(23) = 5.12, p < .001$; implicit learning was not a
significant predictor, $\beta = .16, p > .25$. This suggests that in the stereotype threat condition, participants were relying on their explicit learning point to predict when they learned the categorization rule (see Figure 4), whereas in the control condition, participants relied upon their objective performance to inform this judgment.

### 2.2.8 Motivation

Finally, I examined whether there were differences in motivation to complete the pilot task. All analyses were conducted controlling for the index of objective performance that was deemed most relevant. First, I tested whether participants in the stereotype threat condition would elect to do fewer trials on the ostensible pilot. The ANCOVA (condition: stereotype threat vs. control), controlling for implicit learning point (which was not a significant covariate ($F(1,49) = .159, p = .922$), revealed no significant difference in the number of trials participants elected to do ($M_{threat} = 35.11, SE = 5.86$; $M_{control} = 35.93, SE = 5.86$) $F(1,49) = .10, p = .922$.

I then examined whether implicit and explicit learning points made unique contributions to the number of trials participants elected to do on the pilot task. To test this, I conducted a hierarchical regression in which implicit learning point, explicit learning point, and condition were entered into the first step, the two-way interaction terms were entered into the second step, and the three way interaction was entered into the third step predicting the number of trials participants elected to do on the pilot. This analysis revealed no significant main effects of implicit or explicit learning point ($ps > .1$). There was, however, a significant interaction between implicit learning point and condition $\Delta R^2 = .1, \beta = .49, t(44) = 2.25, p = .029$. To explore this interaction, I conducted a simple slopes analysis. This revealed that in the stereotype threat condition there was a
non-significant positive effect of implicit learning point, $\beta = .36$, $t(44) = 1.24$, $p = .22$. In control condition, there was a marginally significant negative effect of implicit learning point $\beta = - .61$, $t(44) = -1.99$, $p = .052$. This suggests that in the control condition, the fewer trials participants took to learn the categorization rule the more trials they tended to choose on the pilot task, whereas in the stereotype threat condition, participants’ objective performance was not related to the number of trials they elected to do on the pilot.

Next, I tested whether participants in the stereotype threat condition predicted that it would take them longer to learn the categorization rule used in the pilot task. The ANCOVA (condition: stereotype threat vs. control), controlling for implicit learning point (which was a significant covariate ($F(1,48) = 15.97$, $p < .001$), revealed a marginally significant effect such that participants in the stereotype threat condition reported that it would take them somewhat longer to learn the categorization rule ($M_{threat} = 77.2$, SE = 6.03) than did participants in the control condition ($M_{control} = 60.39$, SD = 6.03), $F(1,48) = 3.96$, $p = .052$.

I then conducted a mediational analysis to test the hypothesis that delayed confidence in learning during the categorization task resulted in participants reporting that it would take them longer to learn the categorization rule in the pilot. All mediational analyses were conducted controlling for implicit learning point. As already detailed, participants in the stereotype threat condition were delayed in developing awareness of learning, $\beta = .43$, $t(49) = 2.2$, $p = .033$. Next, I examined the relationship between explicit learning point and the number of trials participants thought it would take them to learn the new categorization rule, while controlling for condition. As hypothesized, explicit
learning point, controlling for condition, was significantly related to the number of trials participants thought it would take them to learn the categorization rule, $\beta = .53$, $t(47) = 3.28$, $p = .002$. To test for mediation, I computed a standardized indirect effect and examined the 95% confidence interval around the indirect effect. The confidence interval for the standardized indirect effect was resampled using the percentile bootstrap (see Biesanz, Falk, & Savalei, 2011). The standardized indirect effect of condition through explicit learning was significant at .22, CI$_{95} = [0.2, 0.49]$. Thus, stereotype threat led to an average .22 standard deviation increase in the number of trials participants thought it would take them to learn the new categorization rule through delayed confidence in learning.
Chapter 3: Conclusion

The goal of this research was to examine how being the target of a negative stereotype could hinder self-perception. The research was motivated by mounting evidence suggesting that being stigmatized might make it difficult to accurately self-assess ability (Aronson and Inzlicht, 2004; Eccles, 1994; Pajares, 2005; Rittmayer & Beier, 2010; Hyde, 2004; Chui & Klassen, 2010). In this present study, I found evidence suggesting that facing a negative stereotype delayed participants’ ability to develop conscious awareness of learning, even when learning itself was not impaired. This was demonstrated by examining how stereotype threat undermined meta-cognitive judgments during an implicit learning task. Participants under stereotype threat showed a decoupling from their objective experience such that they were delayed in becoming aware of their learning. This delayed awareness did not, however, lead to less calibrated metacognitive judgments overall, and measures of post-task perceptions of learning revealed no overall mean differences between the stereotype threat and control condition. However, participants did show differences in what predicted their later assessment of their performance. For participants under stereotype threat, subjective confidence in performance predicted post-task assessments of learning; in the control condition, objective learning predicted post-task assessments of learning. This is suggestive of participants in the stereotype threat condition having better access to explicit than implicit aspects of their performance experience when making post-task assessments of learning. Finally, participants experiencing stereotype threat believed it would take them longer to learn a categorization rule on a subsequent task. This difference was explained by the degree to which participants’ were delayed in developing awareness of their learning.
This provides some support for the idea that the inability to attend to learning has
downstream consequences for perceptions of ability on future tasks.

At the outset of this study, there were two competing hypotheses that could explain threat effects on self-perceived performance: (1) stereotype threat results in lower confidence which delays awareness of learning and, (2) stereotype threat undermines the ability to monitor objective performance which delays awareness of learning. The analytic strategy I adopted looked to test these two potential hypotheses. Based on my analyses, the data more strongly support the conclusion that stereotype threat resulted in some inability to monitor performance, rather than a general underestimation or underconfidence in performance. By examining measures of confidence both during and after the categorization task, I did not find evidence of lower levels of confidence in the stereotype threat condition. However, stereotype threat did delay the ability to become aware of objective performance. That is, in the stereotype threat condition, participants reported consistent explicit awareness of learning about 41 trials after they had displayed consistent evidence of implicit learning. This suggests that participants in the stereotype threat condition were having greater difficulty monitoring their objective performance.

However, a strong conclusion that threat leads to a miscalibration of self-perception is tempered by the fact that analysis of gamma correlations did not yield evidence that stereotype threat results in less accurate perceptions of learning. Gamma correlations provide an index of meta-cognitive accuracy (Nelson, 1994). That is, high gamma correlations are indicative of accurate monitoring of performance. Thus, an inability to monitor performance should have resulted in lower gamma correlations. Averaging across the entire task revealed no condition differences in gamma correlations.
Furthermore, in both conditions, participants evidenced a gamma correlation that was significantly different from zero, indicating that their confidence ratings did track their actual performance. This was also true when examining the gamma correlations before participants became consistently confident in their performance.

These differing findings may be the result of gamma correlations being a less sensitive index of monitoring than the explicit learning criterion. Recent research has pointed to the insensitivity of gamma correlations for assessing the accuracy of metacognitive judgments (Jang, Wallsten, Huber, 2012). New indices are being developed that could be applied to the present research (Jang, Wallsten, Huber, 2012). The explicit learning criterion was highly sensitive to minor fluctuations in confidence. This could also explain why there was not a significant block by condition interaction when examining participants’ FOK ratings across the course of the categorization task. In sum, this research provides some preliminary evidence that stereotype threat makes it difficult to recognize your own competence by undermining the ability to monitor performance. Future research is needed to replicate and strengthen this pattern.

The above findings tell a somewhat unclear story; however, the analyses examining participants’ post-task perceptions of performance offer a stronger pattern of results that provide insight into how stereotype threat affects representations of learning. Participants experiencing stereotype threat appeared to have better access to the explicit aspects of their performance when reflecting upon their learning experience. Furthermore, participants’ relied upon these explicit aspects of learning when making judgments about their future abilities. These findings point to how being the target of a negative stereotype can result in disassociations between explicit and implicit processes.
As already outlined, these disassociations could be result of attentional biases that make it difficult to attend to useful information; or motivated processes that result in useful information being disregarded. Future research should examine whether these inconsistencies result in tension states that give rise to feelings of ambivalence towards learning experiences (Rydel, McConnell, & Mackie, 2008).

This research makes a number of novel contributions to the stereotype threat literature. Schmader and colleagues’ (2008) integrated process model proposes that performance tasks that are not dependent on working memory should not be undermined by stereotype threat. This finding demonstrates that learning which is not dependent upon attentional resources might not be hindered by stereotype threat. This finding provides preliminary support for the idea that learning environments that employ implicit tasks could allow stigmatized students to overcome stereotype threat. Educators might look to implicit learning tasks to teach material. In fact, this has already started to happen. In a recent study, researchers showed that teaching math and science material using implicit learning tasks resulted in improved performance for both 7th grade students and high school sophomores (Kellman, Masey, & Son, 2009; Kellman & Masey, 2010). Employing implicit learning tasks in an educational environment could help in creating a learning environment that more accurately represents stigmatized individuals’ true performance, even if they do not recognize it.

The central goal of this work was to contribute to our understanding of the development of inaccuracies between perceptions of performance and actual performance in stigmatized individuals. There is now a growing body of evidence suggesting that stigmatized individuals have inaccurate self-assessments (Eccles, 1994; Pajares, 2005;
Rittmayer & Beier, 2010; Hyde, 2004; Chui & Klassen, 2010). For example, Aronson and Inzlicht (2004) demonstrated that African American students were miscalibrated in their performance estimates when completing a test. Interestingly, however, minority students in their study overestimated their performance relative to how they actually did. However, in this study the researchers did not manipulate stereotype threat. The present research is the first to manipulate stereotype threat and examine the cognitive processes underlying deficits in self-perception. The current work points to how situational cues in an environment can elicit a series of attentional and motivational biases that undermine self-perception for stigmatized individuals. It offers preliminary support for the idea that inaccurate self-perception might be the result of attentional deficits that mean stigmatized individuals have less access to their objective experience during a performance situation. Furthermore, this research shows that difficulties in attending to experience during a learning situation has consequences for subsequent decision-making. The exact mechanism behind this process is still largely unexplored, but will be examined in future research. Understanding the mechanisms that underlie these effects will provide insight into ways to reduce them. Armed with this knowledge, researcher can design interventions that help stigmatized individuals realize their own potential.

3.1 Limitations and future directions

A limitation of this research is that the mechanism behind the effect has not been established. In this thesis, I presented two potential mechanisms that could lead to inaccurate self-perception: (1) attentional biases and, (2) motivational biases. Both of these mechanisms will need be examined experimentally in future research. To test whether attentional biases can account for the delayed awareness of learning, participants
could complete the category learning task while under cognitive load. The introduction of a cognitive load would test the idea that attentional deficits can result in delayed awareness of learning. Participants would be assigned to one of three conditions: stereotype threat vs. cognitive load vs. control. The stereotype threat and control condition would proceed as they did in the first study. In the cognitive load condition, participants would hear the control instructions, but would also complete a tone counting task during the category learning task. The tone counting task would serve as a cognitive load designed to undermine attention. Performance on the category learning task should not be affected by the tone-counting task (Waldron & Ashby, 2001), but the ability to monitor performance should be hindered (Souchay et al., 2004). This study would provide support for attention as the casual mechanism behind delayed awareness of learning. Future work could also look to measure working memory in order to test it as a mediator of these effects.

The second means by which stereotype threat could hinder awareness of learning is through the introduction of motivational biases. Specifically, I suggested that self-doubt or avoidance motivations could hinder the ability to attend to signals that connote competency. Two experiments could be designed to test whether these motivational states can account for delayed awareness of learning. First, self-doubt could be manipulated prior to completing the category-learning task. Participants would be randomly assigned to one of three conditions: stereotype threat vs. self-doubt prime vs. control. The stereotype threat and control condition would proceed as before. Participants in the self-doubt condition would complete a lexical decision task in which they were primed with self-doubt words prior to completing the category learning task. This should act to
experimentally manipulate the thoughts people are having during the categorization task. Similar methods have been used in past stereotype threat research (Schmader et al., 2009). If the motivational account explains these effects then participants experiencing self-doubt would be delayed in becoming aware of the categorization rule. This research would suggest that thoughts and feelings experienced during stereotype threat (e.g. self-doubt) could hinder the ability to process information that runs contrary to those feelings (i.e. cues connoting competency).

Past research has shown that stereotype threat cues an avoidance orientation (Seibt & Forster, 2004). As previously argued, this motivational state could raise the processing threshold for positive information (i.e. cues connoting competency), resulting in delayed awareness of learning. This could be tested by manipulating participants’ motivational orientation during the category learning task. Past research has demonstrated that motor actions can elicit approach/avoid orientations. For instance, pressing the palm of the hand down on a table thereby activating the extensor muscle, activates an avoidance motivation (Neumann & Strack, 2000). This method could be used to examine how an avoidance motivation makes it difficult to recognize competent performance. Participants could complete the categorization task while pressing down on the table with their non-dominant hand. If approach/avoid orientations can account for these effects then awareness of learning should be undermined by this manipulation. This study would allow me to test the idea that an avoidance orientation can delay the awareness of learning by raising the processing threshold for information that signals competency.
A second limitation of this work is that I did not identify how inaccurate self-assessments could affect subsequent self-views. Future work should look to further explore whether delayed learning assessments results in the development of an inaccurate sense of self. The self-concept is an organizing schema that determines the attitudes, behaviors, and emotions that people express in a given domain and how people process information about themselves (Kunda, 1999). The inability to develop awareness of learning could result in a person’s self-concept related to performance (e.g. efficacy beliefs) being inaccurate (Aronson & Inzlicht, 2004). For instance, a situation in which awareness of learning is delayed might result in the underestimation of ability. Underestimation of ability has been shown to have consequences for performance. For example, researchers have demonstrated that people with overly positive self-beliefs will attempt tasks that they have not yet mastered; whereas, negative self-beliefs result in overly cautious behavior and reduced performance (Bandura, 1997). Thus, inaccurate representations of learning might have implications for performance such that people might attempt less and accomplish less because they feel negatively about their skills and abilities.

Thus, future research might examine how the self-concept could be shaped by the inability to recognize competent performance. For example, participants could do the category learning task under stereotype threat or control conditions and then complete an implicit associations test (IAT) designed to assess the association between self-traits and spatial ability. The inability to develop awareness of learning might result in participants being less identified with spatial ability after the learning task. Another tenable hypothesis is that post task self-concept might not differ by condition, but there could be
differences in the information participants use to inform their self-concept. In the control condition, participants’ objective performance on the categorization task might predict their identification with spatial ability; whereas, in the stereotype threat condition, participants’ subjective perceptions of performance might predict their identification with spatial ability. This study would speak to how a person’s working self-concept is modified or informed by perceptions of performance.

Self-assessments of learning have been shown to have implications for how people allocate study time (Metcalfe & Kornell, 2005). For example, Metcalfe and Kornell’s (2005) region of proximal learning theory suggests that individuals first eliminate already learned items and then choose to study the easiest unlearned items, before moving to the more difficult ones. Research has demonstrated that when people are prevented from studying using meta-cognitively guided strategies, they show reduced learning (Thiede et al., 2003; Kornell & Metcalfe, 2006). The research presented in this thesis suggests that stereotype threat might make it difficult to assess which items have already been learned. Thus, stereotyped students could be at a disadvantage when studying because they are unable to engage in meta-cognitively guided study. This could lead to unsystematic study habits resulting in reduced learning and performance. For instance, stigmatized individuals might perseverate in learning easier already mastered items because they feel uncertain that they truly know the material.

This could be tested experimentally. Participants could complete a version of the category learning task in which they learn multiple categorization rules over the course of the task. Participants would be instructed to move onto a new category rule only when they felt that they had learned the current categorization rule. If participants in the
stereotype threat condition are hindered in meta-cognitively guiding their study, they should perseverate on category rules long after reaching criterion for implicit learning. This would result in participants in the threat condition learning fewer categorization rules over the course of the task. Conducting this study would test whether inaccurate self-assessments of learning have consequences for self-directed study habits.

In conclusion, the present research provides evidence that stigmatized students are somewhat impaired in developing a subjective awareness of what they have learned. In a laboratory experiment, I demonstrated that stereotype threat delays women's ability to develop conscious awareness of implicit learning, even when learning has taken place. Furthermore, this inaccurate awareness of learning predicted participants' post-task perceptions of performance and had downstream consequences for judgments of performance on future tasks. It is my hope that this study will result in a line of research that can be used to outline the mechanisms by which stigmatization results in inaccurate self-knowledge. The resulting findings could inform interventions that foster learning environments that allow stigmatized individuals to recognize their true potential. If learning environments are improved then more minority students should enter into and remain in institutions of higher education. This itself would help reduce the threat of being judged in light of a negative stereotype. Once the threat of being judged in light of a stereotype is reduced, the number of successful minority student role models will increase. An increase in these role models will provide another means for stigmatized individuals to overcome stereotype threat (McIntyre et al., 2005; McIntyre, Paulson, & Lord, 2003). Further, as more minority students enter into fields where they have
traditionally been underrepresented, their very presence will reduce avoidance and underperformance for others.
Figure 1. Performance across the category learning task.
Figure 2. Feeling of knowing ratings (adjusted for performance) across the categorization task.
Figure 3. Number of trials to reach the criterion for explicit learning, controlling for implicit learning point.
Figure 4. Implicit learning point and explicit learning point predicting post task estimates of learning.

**Control**

- Implicit learning (trials) $\beta = .786^{***}$
- Feeling of knowing (trials) $\beta = -.106$
- Post task estimate of learning

**Threat**

- Implicit (trials) $\beta = .162$
- Feeling of knowing (trials) $\beta = .734^{***}$
- Post task estimate of learning

$^{***}p<.001$
Figure 5. Delayed confidence mediating stereotype threat induced under-confidence in future performance.

Condition:
Control = 0
Threat = 1

Delay in confidence

-.21*

.53**

No. of trials to learn new rule

.25

*<.05, **<.01
References


Appendix: Sample Stimuli