2018

Self-efficacy - Performance Discrepancies: Examining How Over- and Underestimations of Ability Progress Over Time

Kent Cooper Etherton
Wright State University

SELF-EFFICACY – PERFORMANCE DISCREPANCIES: EXAMINING HOW OVER- AND UNDERESTIMATIONS OF ABILITY PROGRESS OVER TIME

A thesis submitted in partial fulfillment of the

requirements for the degree of

Master of Science

By

KENT COOPER ETHERTON

B.S., Purdue University, 2015

2018

Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

OCTOBER 31, 2018

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Kent Cooper Etherton ENTITLED Self-efficacy – Performance Discrepancies: Examining How Over- and Underestimations of Ability Progress Over Time BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

Debra Steele-Johnson, Ph.D.
Thesis Director

Scott Watamaniuk, Ph.D.
Graduate Program Director

Debra Steele-Johnson, Ph.D.
Chair, Department of Psychology

Committee on Final Examination

Nathan Bowling, Ph.D.

David LaHuis, Ph.D.

Barry Milligan, Ph.D.
Interim Dean of the Graduate School

ABSTRACT

Etherton, Kent Cooper, M. S. Department of Psychology, Wright State University, 2018. Self-efficacy – Performance Discrepancies: Examining How Over- and Underestimations of Ability Progress Over Time.

The current study examined how over- and underestimations of ability progress with increasing experience completing a task. Prior research has demonstrated inconsistent effects when investigating the relationship between self-efficacy and performance at the within-person level of analysis, often theorizing distinct effects of over- versus underestimating one's ability level. Thus, the current study investigated the discrepancy between self-efficacy, one's belief in their capability to accomplish some task, and actual performance levels. The current study replicated findings that self-efficacy converges on performance over time and extended prior research by demonstrating the rate of convergence might be affected by the size of initial discrepancy and generalized self-efficacy. Further, in a second study meant to compare self-efficacy – performance discrepancies (SPDs) and goal – performance discrepancies (GPDs), we demonstrated that reporting one's self-efficacy before each trial led to better performance than reporting goal level before each trial.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

As jobs are requiring self-guided skill acquisition more frequently, it is important to study the self-regulatory mechanisms people use during the process of acquiring new skills. One such mechanism is self-efficacy, the belief in one's ability to complete a task successfully (Bandura, 1977). At the within-person level of analysis, many researchers have found inconsistent effects of self-efficacy on performance (Feltz, Chow, & Hepler, 2008; Schmidt & DeShon, 2010; Seo & Ilies, 2009; Vancouver & Kendall, 2006; Vancouver, Thompson, Tischner, & Putka, 2002; Vancouver, Thompson, & Williams, 2001; Yeo & Neal, 2006). However, few have examined the "accuracy" of self-efficacy judgments and the extent to which self-efficacy judgments become more accurate (i.e., more closely match performance) with task experience. No one has examined the degree of over- versus underestimation as a distinct variable in the self-efficacy – performance relationship although related research has examined the effects of goal – performance discrepancies and past-performance – performance discrepancies (Kernan & Lord, 1990; Vance & Collela, 1990). Thus, there is a key variable missing in the self-efficacy literature: self-efficacy – performance discrepancy (SPD), or the degree to which one is over- or underestimating his/her own ability and factors affecting change in SPDs. Thus, the purpose of my study was to (a) introduce the concept of SPD to address the role of accuracy in self-efficacy judgements in the skill acquisition process, (b) examine the rate of SPD reduction during skill acquisition, and (c) examine whether SPDs have effects distinct from goal – performance discrepancies (GPDs) or past performance – performance discrepancies (PPDs). We examined the first two issues in Study 1 and the third issue in Study 2.

In the subsequent sections, I address research that provides the foundation for my research purposes and predictions. I begin by addressing skill acquisition and self-regulation

(including goal setting and self-efficacy). Next, I address definitions of and research relating to perceived discrepancies and present my predictions.

Skill Acquisition

Increasingly, the workplace requires people to be able to acquire new skills quickly and effectively. Many jobs require employees to tackle novel challenges frequently, whether it is figuring out a new filing system in the office or learning the steps to submitting paperwork through the proper channels. To accommodate this need for quick learning, it is important to understand the process through which people acquire new skills.

There are two major models of skill acquisition. The first was a three-stage model proposed by Fitts (1964) and later revised by Fitts and Posner (1967). The model was focused on perceptual-motor tasks and how people develop motor skills. The skills Fitts and Posner (1967) tested were highly consistent movements or actions. Although focused on consistent perceptual-motor tasks, the model still has relevance to cognitive tasks.

The cognitive stage, the first stage, involves encoding enough information about task-related behaviors to be able to complete the task. The new behavior consists of multiple previous habits that are organized to achieve the needs of the new behavior. In this stage, the learner attends to cues when developing the new behavior that create "If…then…" procedures, meaning people learn the proper response to a given cue. The same cues will not be attended to consciously later as the behavior becomes more automatic. When learning a new behavior, people must recognize and enact the necessary learned behaviors to accomplish the task. This mental process is considered the cognitive stage of skill acquisition (Fitts & Posner, 1967).

The associative stage, the second stage, is when skill performance is improved as mistakes are identified and eliminated. The second stage involves a degree of associative learning, in which people begin to implement the knowledge they have learned. As new patterns

begin to emerge, errors are gradually eliminated. The associative stage can last for various lengths of time, depending on the skill complexity. The amount of practice needed during this stage varies as well, but in general, frequent repetition of the skill with appropriate rest periods is optimal (Fitts & Posner, 1967).

The autonomous stage, the third stage, refers to when performance of the skill improves, and the behaviors become more automatic. The process begins to be automatized once people gain experience using the new skill and no longer need to think about it. In this stage, people do not have to retrieve bits of knowledge on how to utilize the new skill because they now have formed a heuristic, a single procedure (Fitts, 1964). Performing a task requires less mental processing, and people can learn new skills while exercising the older ones. The speed and efficiency of the skill continue to increase during this autonomous stage. Also, the acquired skills are easier to automate when the tasks are predictable and consistent. By the end of the autonomous stage, people have coalesced multiple units of information into fewer units, allowing for faster, more efficient action (Fitts & Posner, 1967).

The second major model of skill acquisition was developed by Anderson (1982), building upon Fitts' and Posner's theory by labelling the stages differently and by describing how the transition stage works. Anderson's (1982) theory of skill acquisition is distinct from Fitts and Posner's (1967) theory because Anderson's (1982) model addresses how declarative knowledge transfers to procedural knowledge through the creation and consolidation of various mental productions. Anderson (1982) identified the stages of skill acquisition as (1) application of declarative knowledge, (2) compilation of declarative knowledge to create procedural knowledge, and (3) continual refinement of the procedural knowledge. Anderson (1982) divided compilation, the second stage, into two subprocesses: composition, in which sequences of mental productions are collapsed into a single production, and proceduralization, in which declarative

information no longer needs to be retrieved into working memory. Anderson's (1982) theory differs from Fitts and Posner's (1967) theory by focusing on the compilation stage to explain how cognitive information is transferred to automatic behaviors.

Within the compilation stage, Anderson (1982) described a subprocess called composition in which the purpose is to combine multiple independent productions to build single, complex productions. The various simple productions require declarative knowledge to be properly assembled in the correct order to execute complex behaviors. Once the simple productions are combined into a single complex production, declarative knowledge no longer must be attended to when executing the complex behavior, indicating the successful creation of procedural knowledge. When retrieved, procedural knowledge does not require as much cognitive attention/effort because the single complex production makes the behavior more automatic. However, there is a limit to how complex/large productions can become. Working memory must be able to handle the numerous cognitive demands of the complex productions. That is, if the multiple simple productions require too much working memory to perform, it will be difficult to compile them into a single, more complex production.

The other subprocess within Anderson's (1982) composition stage is proceduralization. Proceduralization addresses the composition issue of working memory capacity by reducing the demand on working memory for executing productions. Proceduralization is the process by which long-term memory information no longer needs to be retrieved into working memory to execute a production. Long term memory retrieval becomes built into the production itself. Declarative knowledge no longer needs to be retrieved in working memory because the person has gained procedural knowledge and accesses long term memory directly. Both compilation and proceduralization address the ways in which declarative knowledge is transferred to procedural knowledge.

Various factors affect the skill acquisition process. In the cognitive/declarative stage, the first stage, cognitive ability plays a significant role. The more cognitive ability people have, the more effective they are at acquiring knowledge, subsequently improving their performance (Schmidt, Hunter, & Outerbridge, 1986). Also, motivation is critical for efficient skill acquisition, as it indicates how much time and energy will be exerted on a given task. Exerting time and energy is required when practicing a new skill, and a lack of motivation might delay the automatization of a behavior. Additionally, research has found that self-efficacy is a valid predictor of performance early in the skill acquisition process (Mitchell et al., 1994). However, near the end of the skill acquisition process, goals become the better predictor of performance.

Automatization of new behaviors is rendered impossible for inconsistent tasks. In the skill acquisition literature, automatization is considered possible only after an extreme amount of repetition of the exact same behavior. Assuming the task is consistent, Anderson (1982) mentioned that it would take hundreds of hours for a person to practice any cognitive skill to automaticity.

As mentioned earlier, acquired skills are easier to automate when tasks are predictable and consistent (Fitts & Posner, 1967). Schneider and Shiffrin (1977) proposed a model of attention within the context of visual search and detection. In their model, they described and found support for two qualitatively different mechanisms they called automatic processing and controlled processing. For automatic processing to occur, Schneider and Shiffrin (1977) found that stimuli needed to be consistently mapped to responses. If the stimuli were not consistently mapped, it would be more difficult for a person to develop a strategy and automatize the process. Some tasks are difficult, even impossible, to automatize due to their ever-changing task components. When tasks are too inconsistent to automatize, people instead develop habits through repetitive learning.

In the training literature, researchers have used the term ‘overlearning’ to address skill acquisition to a high level of proficiency on complex and inconsistent tasks. Overlearning is intentional practice beyond a set performance criterion. When tasks are impossible to automatize, overlearning can reinforce behaviors toward eventual expertise by repeating the behaviors beyond the point of minimum acceptable performance. Empirically, Krueger (1930) was one of the first to study the effects of overlearning. Participants repeated a task until they reached a set minimum threshold of performance and then continued practicing the task in additional trials. The additional practice led to increased levels of retention. The goal of overlearning is to increase retention of information and minimize memory decay. Meta-analytic findings indicated a moderate effect for overlearning on retention, and this effect was moderated by the degree of overlearning, type of task, and length of retention period (Driskell, Willis, & Copper, 1992). Repeating the mental processes involved in a complex task strengthens the habit created by repetitive practice.

In the current study, I will examine the role of self-regulatory behaviors in the skill acquisition process. Primarily, I will focus on the effect motivational variables have on effort and subsequent performance in the early stages of skill acquisition (i.e., when declarative knowledge is gained and applied in practice and simple productions are formed).

Self-regulation

Self-regulation affects how quickly and effectively people learn information and acquire new skills. Broadly defined, self-regulation is the motivational process that guides the allocation of time to and effort applied toward attaining a goal (e.g., Bandura, 1991; Carver & Scheier, 1982). An unmotivated person is unlikely to exert as much effort to either encode or retrieve relevant declarative knowledge. Effort is used in both the cognitive and associative stages of skill acquisition because declarative knowledge is being encoded and translated into procedural

knowledge through practice. Effort is required to learn and apply declarative knowledge when practicing a skill, and the effectiveness of that effort depends on how well that person regulates his or her behavior.

Two theories describing self-regulatory processes are Control Theory and Social Cognitive Theory. Control Theory is a cybernetic theory of self-regulated behavior that originated in the engineering field and was modified for human behavior by Carver and Scheier (1982). The process underlying Control Theory is a negative feedback loop, in which people are trying to reduce perceived discrepancies between current states and desired states. When people set goals and notice discrepancies between their self-set goals and performance, they either exert effort toward reducing that discrepancy or simply relinquish the goal. Frequently, researchers reference Control Theory when discussing why certain perceived discrepancies influence motivational variables and effort.

Control Theory has three main components: standards, monitoring, and operation (Carver & Scheier, 1982). Standards are ideals or goals a person holds. Monitoring is the process by which a person compares his actual state to his standards. This is the stage in which perceived discrepancies manifest. Operation is the process through which effort is exerted to reduce perceived discrepancies between the person's actual state and his/her standards. These three components are considered the primary mechanisms through which people reduce perceived discrepancies and self-regulate their behavior.

A hierarchy of goals influences the standards used in Control Theory's negative feedback loop (Carver & Scheier, 1982). Powers (1973) described goal setting as having a hierarchy of 'quantities' which contribute to the attainment of higher-order goals. For example, if someone has a goal to exercise more often, there are a near-infinite number of lower-order 'quantities' which contribute to attaining such a goal, including 'developing a habit of exercising', 'going

outside', or even 'the activation of leg muscles involved in movement'. Higher-order goals cannot be attained without attainment of the lower-order quantities.

In another self-regulatory theory, Social Cognitive Theory, Bandura (1986) suggested that human behavior is regulated by the ongoing exercise of self-influence (Bandura, 1986). Within self-regulation, there are three principal subfunctions: self-monitoring of one's behaviors, judgement of one's behavior compared to personal standards and environmental circumstances, and affective self-reactions (Bandura, 1991). These three subfunctions constitute the structure of this self-regulatory system. Functionally, this system is influenced by self-efficacy, the belief in one's ability to execute a particular behavior successfully (Bandura, 1977). Bandura (1991) claimed self-efficacy is the most central mechanism of human agency because it influences the subfunctions of the self-regulatory system.

The self-monitoring subfunction addresses the need for self-reflection in the self-regulatory process. It would be impossible for a person to regulate his behavior if he did not pay attention to it. Bandura (1991) described self-monitoring as having a self-diagnostic function, allowing people to notice patterns about their own behaviors, and a self-motivating function, allowing people to set realistic goals for themselves as well as monitor their progress toward accomplishing goals. Self-monitoring can be influenced by many factors such as preexisting cognitive structures, self-beliefs, perception of one's functioning, how performance information is organized for memory encoding, and mood. The likelihood of self-monitoring causing change in an individual depends on the temporal proximity of the self-monitoring to behavior and the informativeness of the performance feedback (Bandura, 1991).

The judgmental subfunction involves a comparison between self-observed performance levels and various standards. People use this subfunction to determine whether they have achieved their desired performance levels. People acquire information regarding their

performance level from the self-monitoring subfunction whereas standards can be acquired through various sources depending on the person's self-monitoring orientation (Snyder, 1987). Standards are obtained from an interaction between self-generated and external sources of influence. This interaction includes reflective processing of multiple sources of direct and vicarious influence such as the tutelage of others, reactions of others, and social referential comparisons. Once standards are established, a person has a reference point with which she can compare her performance level. However, a person's valuation of the activity can determine whether the judgement of performance will be used to exert effort or simply cease performing the activity. A person who does not value the activity being performed is less likely to spend time and cognitive resources to reduce the perceived performance discrepancy.

The affective self-reaction subfunction is the emotional response to achieving certain performance levels. Self-reactions are considered the consequences of performance that cue self-regulation. Self-reactions are the outcomes of the self-monitoring and judgement subfunctions. A person needs to monitor her performance and compare it to her given standards before knowing how she should feel about her performance level. Also, self-reactions depend on the perceived performance determinants, or what factors led to a person's success or failure. If the person succeeds due to external determinants, he will be less likely to derive self-satisfaction from his accomplishment. When a positive self-reaction is anticipated from achieving a given goal, a person becomes more motivated to accomplish said goal. People pursue activities that produce positive self-reactions and avoid activities that produce negative self-reactions (Bandura, 1991).

Monitoring, judgement, and affective self-reactions are all influenced by self-efficacy, the belief in one's ability to perform a task or more specifically to execute a particular behavior successfully (Bandura, 1977). When a person monitors his performance level, the information

gained likely will influence his self-efficacy. Conversely, self-efficacy influences performance through goal setting. Those with high self-efficacy tend to set higher goals for themselves. Also, the valuation aspect of the judgement subfunction, how much people value the task they are performing, is affected by self-efficacy. People show more interest in activities they believe themselves to be good at (Bandura & Schunk, 1981). Due to its wide applicability within the self-regulatory process, Bandura (1991) considered self-efficacy to be the most central mechanism of personal agency.

Bandura (1991) claimed self-efficacy is significant in both discrepancy production and discrepancy reduction systems. Discrepancy production involves goal-setting, in which self-efficacy affects the goal level a person sets for him- or herself. Discrepancy reduction is the process of working toward a set goal, reducing the perceived discrepancy between current states and ideal/goal states. To reduce perceived discrepancies, people use self-efficacy to determine how much effort is needed to achieve their set goal.

Social Cognitive Theory described an additional model called reciprocal determinism. This model stated that cognitive, behavioral, and environmental factors all influence each other when predicting psychological functioning (Bandura, 1978). These three main factors often affect each other in some way. Social Cognitive Theory focused on the effect that social/environmental factors have on cognitive regulation of behavior. From this perspective, it is easy to see how self-efficacy, a cognitive factor, might influence a person's behavior and how a person's behavior might influence his/her self-efficacy.

Both people in the workforce and students engage in self-regulatory behaviors. People in the workforce constantly engage in self-regulatory behaviors that enable them to work effectively. For example, managers self-regulate their feedback-seeking process which has significant effects on their fellow workers' (i.e., supervisors, peers, subordinates) perceptions of

their managerial effectiveness (Ashford & Tsui, 1991). Setting goals in the work setting can lead to an increase in productivity and a decrease in absenteeism (Latham & Kinne, 1974).

Also, self-regulation is a critical topic for educational environments. Zimmerman (1990) defined self-regulated learning as the process of acquiring information or skills that involve agency, purpose, and instrumentality perceptions by learners. Everyone experiences this process sometime during the first ~18 years of life and would benefit from a scientific understanding of the process. Research has found that a person's beliefs in his/her self-regulated learning indirectly influences academic achievement through academic self-efficacy and self-set goals (Zimmerman, Bandura, & Martinez-Pons, 1992). Understanding the process of self-regulation is beneficial to a variety of fields, which is why it is important to study the components of self-regulation. Two primary components of self-regulation are goal setting and self-efficacy, each enabling one to self-regulate their behavior.

Goal setting. When researchers first began to examine goals, one prominent psychological construct was level of aspiration. Dembo (1931) found that when required goals are too difficult, people set intermediate goals (as referenced in Ricciuti, 1951). She called these intermediate goals the momentary level of aspiration. The first major study of level of aspiration was done by Hoppe (1930), who examined the nature of level of aspiration, finding it to fluctuate in response to perceived success and/or failure (as referenced in Ricciuti, 1951). Often, researchers have examined how people perceive success and failure when they compare their performance levels to their levels of aspiration (Hoppe, 1930, as referenced in Ricciuiti, 1951; Sears, 1940). Hoppe (1930) conceptualized level of aspiration as an abstract combination of constantly shifting expectations, goals, and demands (as referenced in Ricciuti, 1951). Empirically, level of aspiration was studied as a point of comparison with performance levels.

Later, researchers questioned whether the process of setting goals might engage motivational behavior (Locke, 1968; Locke & Latham, 1990).

Locke (1968) argued that goals themselves motivate people to take action. He believed any theory addressing task-motivation must involve conscious goals and intentions. Locke (1968) proposed that goal difficulty is positively related to performance. Setting more difficult goals leads to better performance. Also, goals should be specific because people with specific hard goals typically outperform those with vague or "do your best" goals. Finally, Locke (1968) argued that behavioral intentions regulate choice behavior. Often, goals and intentions mediate the effects of external incentives. For example, behavioral intentions were found to mediate the effects of money and verbal reinforcement on choice behavior (Locke, 1968).

Locke and Latham extended Locke's (1968) assertions with Goal-Setting Theory, suggesting that goals are regulators of human action (e.g., see Locke, 1968; Locke & Latham, 1990 for reviews). Locke and Latham (1990) classified variables that affect the goal setting process as goal content or goal intensity variables. Goal content variables were related to the outcome of the goal (e.g., goal level and goal specificity). Goal intensity variables consisted of factors like goal commitment and the importance of the goal. Research has found goals that are specific, difficult, and attainable, provided the person is committed to the goals, produce higher levels of performance (e.g., see Locke & Latham, 1990 for a review).

There are four main variables that mediate the goal setting – performance relationship: effort, persistence, direction of attention, and requisite ability (Locke & Latham, 2006). Assuming adequate ability, greater effort is required to accomplish difficult goals, and exerting more effort leads to better performance (Locke & Latham, 1990). Setting specific, challenging goals leads to effort being exerted over a longer period of time, otherwise known as increased persistence (Bavelas & Lee, 1978). When people set goals, they direct their attention toward

accomplishing the goals, which leads to retrieval of relevant task knowledge (Locke & Latham, 1990). Finally, goals prompt people to retrieve relevant task knowledge. If relevant task knowledge is unavailable, such as with a complex, novel task, the person with difficult goals will be motivated to seek new knowledge.

The key moderators of the goal setting – performance relationship are feedback, commitment, task complexity, and situational constraints (Locke & Latham, 2006). As the quality of feedback increases, goal-setting will have a stronger positive relationship with performance. People will be able to more accurately monitor their progress toward accomplishing their goal. Commitment is required for effective goal setting because otherwise people would not be invested nor exert energy toward accomplishing the goal. For complex tasks, it is more difficult to obtain task-relevant information, which means the effect of goal setting on performance is attenuated. Finally, research has found that situational constraints, such as completeness of task information, ease of use of materials, and similarity of the work environment to the training environment weakened the relationship between goal setting and performance (Peters et al., 1982).

Valence and expectancy affect goal commitment and goal choice (Locke & Latham, 1990). A person will not be committed to a goal if he/she does not value it or if he/she does not expect to be able to accomplish it. The valence associated with a goal is determined by several variables such as group norms, normative information, and goal assignment (see Locke & Latham, 1990, for a comprehensive list).

The expectancy associated with a goal is influenced by previous performance, ability, and self-efficacy (see Locke & Latham, 1990 for a comprehensive list). People set goals based on their perceived personal ability reference point. If a person is highly efficacious, she will set higher goals because she believes she can achieve a higher level of performance (Locke &

Latham, 1990). Also, those with high self-efficacy are more likely to be committed to assigned goals and use better task strategies to attain set goals (Locke & Latham, 1990; Seijts & Latham, 2001).

Self-efficacy. From its conception in the mid-1970s until around 2000, self-efficacy was analyzed primarily at the between-person level of analysis and had a consistent positive relationship with performance. Bandura (1977) claimed when individuals have a high level of self-efficacy, they are more likely to succeed at a given task. This means that the stronger individuals' beliefs that they can accomplish some goal, or complete some task, the more likely they will perform well. Conversely, if individuals believe they will not be able to succeed at a given task, they are more likely to perform poorly. Stajkovic and Luthans (1998) found an average corrected correlation between self-efficacy and work-related performance of $\rho = .38$ with the relationship being moderated both by task complexity and type of study setting. Often, researchers incorrectly interpreted this relationship as meaning that those with high self-efficacy perform better *because* of their self-efficacy.

Recent research has examined inferred causal relationships between self-efficacy and performance through empirical testing (e.g., Heggestad & Kanfer, 2005; Sitzmann & Yeo, 2013; Vancouver et al., 2002; Vancouver, Thompson, & Williams, 2001). Given that past performance influences self-efficacy, searching for a causal effect of self-efficacy on performance implies a possible reciprocal relationship. Self-efficacy might affect subsequent performance level, and previous performance informs subsequent self-efficacy.

Since 2000, researchers have searched for evidence of a causal relationship between self-efficacy and performance using a within-person level of analysis and have found inconsistent results. Vancouver, Thompson, and Williams (2001) measured the self-efficacy of participants periodically while they played an analytic game. Vancouver, Thompson, and Williams (2001)

found that participants with higher self-efficacy performed worse than those with lower self-efficacy. Vancouver, Thompson, and Williams (2001) concluded that those with inflated confidence were exerting less effort in the game and consequently were making more errors. Using a within-person design, in two different learning contexts, Richard, Diefendorff, and Martin (2006) found that past performance predicted self-efficacy, but self-efficacy failed to predict subsequent performance. Heggestad and Kanfer (2005) conducted a within-person study in which self-efficacy measures were taken before and after various tasks. They found that after partialing out previous self-efficacy from past performance improvements, self-efficacy was no longer related to subsequent performance. Heggestad and Kanfer (2005) concluded that self-efficacy is likely a consequence and not a cause of performance. There are many studies finding conflicting results regarding self-efficacy's effect on performance from the within-person perspective. Until now, research has not considered the possible role of inaccurate self-efficacy judgements and the resulting perceived discrepancy.

Perceived Discrepancies

An important aspect of self-regulation research is how people create and regulate perceived positive and negative discrepancies. Social Cognitive Theory stated that discrepancy production and discrepancy reduction are both self-regulatory systems through which self-efficacy affects performance (Bandura, 1991). Control Theory contended that people strive to reduce perceived negative discrepancies between current states and desired states (Carver & Scheier, 1982). When people set goals and notice discrepancies between their self-set goals and performance, they might exert effort toward reducing that discrepancy or relinquish the goal.

Researchers have found robust effects when studying goal – performance and past performance – performance discrepancies. Sears (1940) examined the discrepancy between a person's level of aspiration (i.e., goal level) and her actual level of performance on a previous

trial, finding evidence that attitudes of success and failure influence goal levels. Those who had experienced success in a previous trial had less variability in their goal – performance discrepancies than those who had experienced failure (Sears, 1940). Research has found that assigned goals are rejected when the goal – performance discrepancy becomes sufficiently negative (Vance & Collela, 1990). Also, goal – performance discrepancies strongly affect goal priority and resource allocation (Kernan & Lord, 1990). Vance and Collela (1990) noted that goal – performance and past performance – performance discrepancies are used differently, particularly in the goal evaluation process. Researchers have treated these discrepancies as distinct variables, leading to further knowledge on the process by which people use goals and react to performance discrepancies.

Self-efficacy – Performance Discrepancies (SPDs). Largely, research has ignored discrepancies between self-efficacy judgements and performance. Little research has examined the effect that over- and underestimates of self-efficacy have on subsequent performance. Moores and Chang (2009) tested a model using structural equation modeling at the between-person level of analysis. Overconfidence was measured meta-cognitively as the difference between a person's expected performance and his/her actual performance. After assessing the difference, Moores and Chang (2009) categorized the participants into over-confident versus under-confident groups. They found that self-efficacy had a non-significant relationship with performance across all participants but a significant negative relationship with performance for the over-confident participants. This illustrates the importance of over- and underestimations of ability in the self-efficacy literature, a massive literature that has overlooked this simple consideration.

Also, researchers have examined a construct similar to overestimation: overplacement. Overplacement is a comparative form of overconfidence. For example, people might say they

performed in the 90th percentile when they actually performed at the 50th. Overplacement is a difference of percentile scores instead of raw scores. Emich (2014) found that positive affect reduces overplacement but has no effect on overestimation. Overplacement falls under the 'overconfidence' umbrella but is considered distinct from overestimations. Recent research has distinguished between overestimation, overplacement, and overconfidence, but older research often has used the terms interchangeably. Over- and under-placement is comparative to others whereas over/under-estimations are personal judgements.

Sitzmann and Johnson (2012) found that over- and underestimation is related to numerous variables including performance, attrition, effort, commitment, and conscientiousness. For example, the effect of performance on subsequent performance and attrition was moderated by self-assessments of knowledge. That is, Sitzmann and Johnson observed a stronger relationship between current performance and subsequent performance for those participants who underestimated their ability compared to those participants who overestimated their ability. Similarly, effort was more strongly and positively related to performance for participants who underestimated rather than overestimated their ability. Also, overestimations had a greater effect on attrition from goals, compared to underestimations. Additionally, there was a main effect of estimations of ability on performance in that participants who overestimated their ability demonstrated lower performance relative to participants who underestimated their ability. Thus, research has shown that numerous variables are affected by different types of inaccurate self-efficacy judgements, i.e., over- and underestimations of ability.

Researchers have used various methods to calculate over- and underestimations of ability. Moores and Chang (2009) created difference scores between self-efficacy and performance but then treated over- and underconfidence as distinct categorical variables. These researchers assigned participants into either an 'overconfident' group or an 'underconfident' group,

depending on whether the participant's expected performance was more than one percent above or below his performance level. Other research has tested over- and underconfidence as the interaction between self-efficacy and performance level while predicting subsequent performance (Sitzmann & Johnson, 2012). Using continuous discrepancy scores, the goal setting literature has found substantial effects using goal – performance discrepancies and past-performance – performance discrepancies to predict future outcomes such as performance (Kernan & Lord, 1990; Vance & Colella, 1990).

In the self-efficacy literature, there is a need for a concrete variable that can explain potential differential relationships between self-efficacy and performance at the within-person level of analysis. Research has not yet examined the magnitude of or change in a person's over- or underconfidence over time. To examine these effects, I propose a new construct, self-efficacy – performance discrepancy (SPD), which might clarify the examination of over- and underestimations of future performance.

A key variable missing in the self-efficacy literature is the degree to which one has over- or underestimated his/her capability to complete a task, or his/her self-efficacy – performance discrepancy (SPD). An SPD is the difference between a person's anticipated level of performance and his/her actual level of performance. That is, an SPD is the difference between a prior expectation and the subsequent performance level.

If a person is provided with useful feedback on a task he is assigned to complete, he will likely reduce the absolute value of SPD over time. The primary tenet of self-regulation is discrepancy reduction. Being a key self-regulatory variable, self-efficacy likely will adhere to the familiar pattern of self-regulatory discrepancy reduction. Research has found that the best predictor of self-efficacy is past performance (Atkinson, 1964; Heggestad & Kanfer, 2005; Sitzmann & Yeo, 2013), suggesting that people's self-efficacy assessments will converge with

their actual performance levels given enough time and experience. When making self-efficacy judgments, people use as much data as is available, so with increasing amounts of experience with a task, a person will have more information with which to derive an accurate expectation. With increasing task experience (trials of task practice) and useful feedback, the absolute value of SPDs likely will decrease.

Hypothesis 1: Task experience will be negatively related to |SPD|s such that |SPD|s will decrease over time.

Moreover, people likely will decrease their SPDs at different rates, and this difference might be due in part to people's responses to negative vs positive feedback. For example, Kernis and Johnson (1990) found a stronger relationship between self-appraisal change and emotional reaction following negative feedback as opposed to positive feedback. When people receive negative feedback, those who experience threat to their self-perceptions engage in a different emotional process and are more likely to reassess their self-efficacy than people whose self-perceptions are not threatened (Baumeister, Smart, & Boden, 1996, Niemann et al., 2014). Therefore, people who overestimate their own ability level (i.e., negative feedback) should be more likely to reassess their self-efficacy and reduce perceived discrepancies. This process likely will influence their confidence levels in subsequent trials such that positive SPD_1 will lead to a stronger negative effect of Trial on subsequent |SPD|s.

Hypothesis 2: Trial will interact with the sign of initial SPDs when predicting future |SPD|s such that those initially overestimating their ability level will reduce their |SPD|s more quickly over time.

Additionally, the magnitude of an initial SPD might affect the size of subsequent SPDs. Control theory and Social Cognitive Theory posit that a behavior is performed if there is a perceived discrepancy between the current state and the goal state/reference value (Bandura,

1989; Carver & Scheier, 1982). Then, the individual engages the behaviors to reduce the perceived discrepancy. Larger discrepancies are easier to notice, which increases the chance of them being perceived, and subsequently reduced through exertion of effort and attentional resources. Similarly, for SPDs, those individuals who produce and perceive a large initial SPD likely will exert additional effort, or simply be more likely to notice the discrepancy, and consequently reduce the SPD to a greater extent than if they exhibit and notice a smaller initial SPD. This different process likely will influence their confidence levels in subsequent trials such that large $|SPD_1|$ will lead to a stronger negative effect of Trial on subsequent |SPD|s.

Hypothesis 3: Trial will interact with the size of initial SPDs when predicting future |SPD|s such that those who initially under-/overestimate their ability level to a larger extent will reduce their |SPD|s more quickly over time.

Further, some people might be more motivated to reduce SPDs than others, and that difference might be due to personality factors. Some people are more likely to reassess their self-efficacy after receiving negative feedback, such as those who are emotionally stable and less prone to egotism (Baumeister, Smart, & Boden, 1996; Ilgen & Davis, 2000; Niemann et al., 2014). Other dispositional traits might influence the likelihood of reassessing their self-efficacy in response to feedback. For example, people who are high in some facets of agreeableness (e.g., Trust in Others, Modesty) might be more likely to reevaluate their inaccurate self-efficacy ratings. Also, other personality factors such as cautiousness (facet of conscientiousness), trait optimism, trait pessimism, and core self-evaluations might affect initial and subsequent self-efficacy ratings although the nature of those effects is unclear.

Research Question: Trial and personality factors will interact when predicting |SPD|s.

Researchers have found significant effects of goal – performance discrepancies (GPDs) and past performance – performance discrepancies (PPDs) on various outcomes (Kernan & Lord,

1990; Vance & Collela, 1990). The use of said discrepancy variables has enabled research to address different and unique questions regarding the effects of goals and prior performance. Also, prior research has found that self-efficacy better predicts performance early in skill acquisition than goal setting (Mitchell et al., 1994). This indicates that self-efficacy captures some unique content in performance, indicating that some unique content within self-efficacy is distinct from goal setting in motivational processes. Further, previous research has found that self-efficacy has a direct effect on performance, even controlling for goal-setting and previous performance (Zimmerman, Bandura, & Martinez-Pons, 1992). If the motivational influence of self-efficacy is indeed distinct from that of goals and previous performance, then self-efficacy – performance discrepancies should account for incremental variance in performance beyond that accounted for by goal – performance and past-performance – performance discrepancies.

H4: SPDs will account for incremental variance in performance beyond the variance accounted for by GPDs and PPDs.

Method

I conducted two studies to examine the construct of self-efficacy – performance discrepancies (SPDs). The first study evaluated the nature of SPDs over a series of trials (i.e., whether SPDs decrease, increase, or remain constant over time). The second study examined distinctions between SPDs and other motivational discrepancy variables, i.e., goal – performance discrepancies (GPDs) and past-performance – performance discrepancies (PPDs).

Study 1: Examining the Nature of SPDs

My first three hypotheses focus on the nature of SPDs. The first hypothesis stated that SPDs will decrease in size over time as people gain more experience with the task. The second and third hypotheses stated that the effect trial has on later SPDs is moderated by the sign (Hypothesis 2) and magnitude (Hypothesis 3) of initial SPDs. Also, I examined a research

question which states that personality characteristics will interact with trial and the size of the initial SPD when predicting subsequent SPDs.

Participants

Participants in this study were students in an introductory psychology course at a mid-sized Midwestern university. Participation in this study counted toward fulfillment of a requirement in the introductory psychology course.

Class Scheduling Task

The class scheduling task is a computerized task developed by Steele-Johnson (personal communication, May 5, 2016) in which the objective is to earn as many points as possible. Points are earned by creating class schedules for as many fictional students as possible while adhering to certain rules. Participants earn 20 points for every successful class schedule and lose 5 points for each violated rule. The 5-point penalty is removed if the violation is corrected. The participants must adhere to 7 rules (e.g., no classes may overlap, the student prefers classes in the AM, this class requires a lab section). Whenever a rule is violated, a message will appear on their screen indicating which rule has been violated. When the participant corrects the rule violation, the error message will disappear, and the 5-point penalty is removed.

The scheduling task display has four windows (see Figure 1). The top left window shows the available courses. To create a schedule, the participant must browse through a set of courses to find courses which satisfy each fictional student's preferences. There are 42 courses that all have multiple sections (~2-6 sections per course) on different days and at different times. The top right window shows a fictional student's schedule. In each fictional student's schedule window, there is an ID number, class time (AM/PM/Evening) preference, day of week (MWF/TR) preference, and 5 empty lines for courses. The empty lines will be filled in as the participant transfers courses from the available courses window (top left) into the fictional

student's schedule window (top right). The bottom left window shows the number of completed schedules ("NO. SCH"), how many points have been earned from completing schedules ("PTS. SCH"), the current number of errors committed ("NO. ERR"), the points deducted because of errors ("PTS.ERR"), the cumulative point total ("CUM"), how much time has passed in the current trial, and rule violation messages. There is enough room in the error window to display every error message simultaneously. Participants can continue creating schedules while error messages are displayed. The bottom right window is a review window, displaying other fictional students' schedules. The review window is intended to allow the participant to view two schedules simultaneously so that the participant does not create an identical schedule for two separate fictional students, which would violate a task rule.

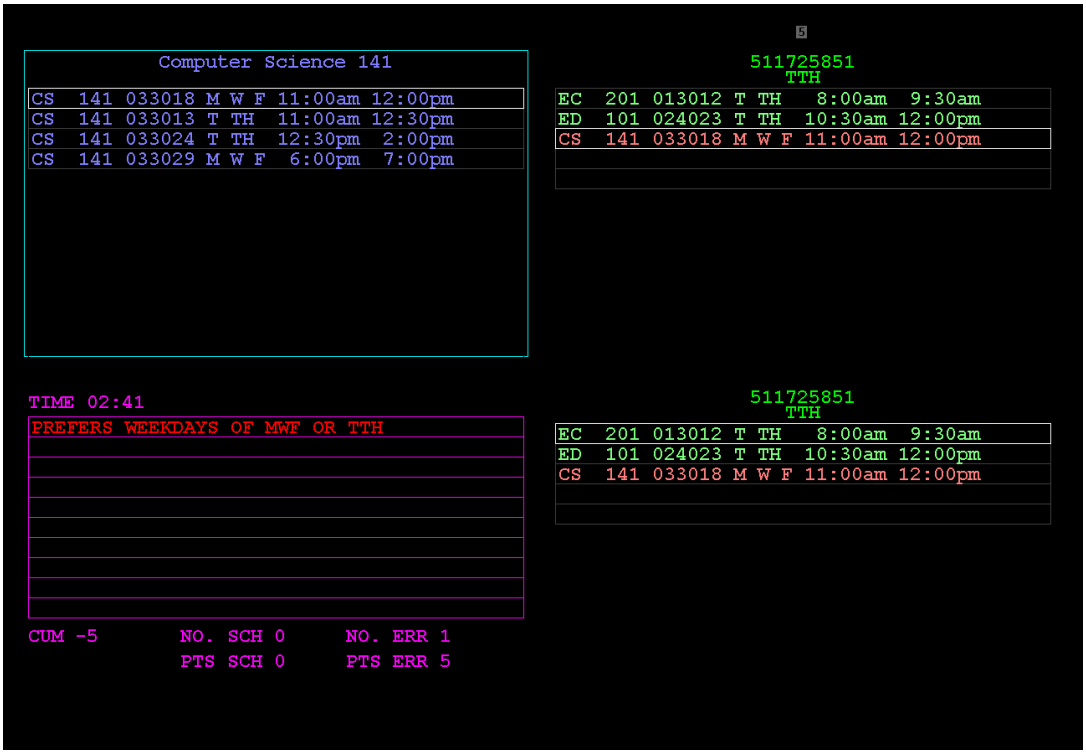

Figure 1. Class Scheduling Task display.

Participants received task instructions describing how to (1) navigate the interface, (2) change which fictional student's schedule is displayed in the top-right and bottom-right

windows, (3) change the course displayed in the available courses window (top-left), and (4) transfer courses from the available courses window to the fictional student's schedule window.

Measures

Demographics. I intended to measure demographics in a survey asking participants about sex, age, year, major, GPA, ethnicity, and handedness (see Appendix A). I intended to ask participants about their handedness because the class scheduling task involves near exclusive use of the right hand. If people are left-handed, they might perform poorly on the task simply by using their non-dominant hand. Unfortunately, due to a programming error, demographic data was not collected for Study 1.

Task-specific self-efficacy. I measured task-specific self-efficacy by revising Bandura's (2006) measure into an 11-item measure of participants' confidence regarding performing at different levels on the subsequent trial. I followed recommendations by Bandura (2006) for revising the self-efficacy measure. For each performance level, participants answered how confident they are in their ability to achieve the given performance level if they were to complete the task again at that current moment. Bandura (2006) mentioned "it is easy for people to imagine themselves to be fully efficacious in some hypothetical future", so it is more appropriate to ask a participant about their self-efficacy at that current moment. Responses are rated on a graphic ratings scale from 1 (Cannot do at all) to 9 (Highly confident can do). Typically, self-efficacy measures with this format are used to calculate an overall efficacy score for the task. In the current study, I needed a single performance level to compare with the actual performance level. For this reason, I decided to use the highest performance level the participant is at least 80% confident they can achieve. The measure is scored by recording at which performance level their confidence ratings reach 8. If a participant did not assign a confidence rating of 8 to any performance level, I used the performance level with the next highest confidence rating (i.e., 9).

If the participant's confidence ratings never exceeded 7, the self-efficacy score was not recorded for that trial. Any participant who responded with the exact same confidence rating for each of the eleven items was excluded from analyses due to insufficient effort responding (unless they respond with all 9s). Any participant who is paying attention would realize that their confidence level for attaining zero schedules should be different from their confidence level of attaining 13 schedules. An example item is "Cumulative point score ≥ 20 (1 Schedule)". See Appendix B for scale items.

Performance. Performance was operationally defined as the participant's cumulative point total from a given trial of the Class Scheduling Task. Participants earn 20 points for every completed schedule and lose 5 points for each error violation. Points deducted for error violations are replaced for each error the participant fixes.

Self-efficacy – Performance Discrepancy (SPD). I obtained the absolute value of the difference between the participant's self-efficacy and their level of performance for each trial, resulting in 8 SPDs.

Personality. I measured personality using the 50-item measure of the Big Five personality factors and 28 additional items from Costa and McCrae's (1992) NEO-PI-R IPIP measure for a total of 78 personality items (International Personality Item Pool, 2013). I added the additional facet items because I expected three facets in particular to be most influential in the current study: trust in others (Agreeableness; 10-items), modesty (Agreeableness; 10-items), and cautiousness (Conscientiousness; 10-items). For the five domains, Costa and McCrae (1992) reported Cronbach's alphas of 0.86 (Extraversion), 0.86 (Emotional Stability), 0.82 (Openness), 0.77 (Agreeableness), and 0.81 (Conscientiousness). For the facets, Costa and McCrae (1992) reported Cronbach's alphas of .82 (Trust in others), .77 (Modesty), and .76 (Cautiousness). Participants were asked to rate all items using a graphic rating scale, ranging from 1 (Very

Inaccurate) to 5 (Very Accurate). Sample items included "make friends easily" (Extraversion), "often feel blue" (Emotional Stability), "have a vivid imagination" (Openness), "respect others" (Agreeableness), and "am always prepared" (Conscientiousness). Item responses were averaged to provide an overall score for each factor. Also, facet scores were calculated by taking the average of items representing each facet. Higher scores indicated higher levels of the corresponding domain or facet of the Big Five. See Appendix C for scale items.

Optimism. I measured dispositional optimism using the 10-item Life Orientation Test-Revised (LOT-R) optimism measure developed by Scheier and Carver (1985) and later revised by Scheier, Carver, and Bridges (1994). Responses were rated on a graphic rating scale from 1 (Strongly Disagree) to 5 (Strongly Agree). The measure was scored by averaging the participant's responses to compute an overall optimism score. The internal consistency reliability for the scale is .78 (Scheier, Carver, & Bridges, 1994). Higher scores on the scale indicated higher levels of optimism. An example item is: "In uncertain times, I usually expect the best". See Appendix D for scale items.

There is some dispute regarding whether the construct of optimism is distinct from pessimism or the two are on opposite poles on a single continuum (Marshall et al., 1992; Scheier, Carver, & Bridges, 1995). Originally, Scheier and Carver (1985) found evidence of a two-factor solution to the Life Orientation Test (LOT) measure of optimism but continued using a single-factor solution. They made the decision because they believed the two-factor solution was due to differences in item wording instead of meaningful item content. When Scheier, Carver, and Bridges (1994) revised the LOT, they tested the factor structure again using both orthogonal and oblique rotations, both providing evidence of a single-factor solution. Despite the single-factor solution, Scheier, Carver, and Bridges (1994) still recommend that subsidiary analyses use a two-factor solution. Because of the findings of Scheier, Carver, and Bridges (1994), I conducted

exploratory factor analyses on the LOT-R and determined the single-factor solution fit best. Therefore, I treated optimism and pessimism as opposite poles of a unidimensional construct.

Core Self-Evaluations. I measured core self-evaluations using the 12-item Core Self-Evaluations Scale (CSES) developed by Judge, Erez, Bono, and Thoresen (2003). Responses were rated on a graphic ratings scale from 1 (Strongly Disagree) to 5 (Strongly Agree). The measure was scored by taking the average of item responses. The internal consistency reliability for the CSES ranged from .81 to .87 (Judge, Erez, Bono, & Thoresen, 2003). Higher scores on the scale indicated higher levels of Core Self-Evaluations. An example item is "When I try, I generally succeed". See Appendix E for scale items.

Trait anxiety. I measured trait anxiety using the 6-item abbreviated version of Spielberger's (1983) State/Trait Anxiety Inventory (STAI) scale developed by Marteau and Bekker (1992). Typically, the STAI measures trait and state anxiety by asking participants to respond to six items given two separate sets of instructions (one set for trait and one set for state). I used only the trait anxiety instructions, which ask participants to respond to the items by indicating how they "generally feel". Responses were rated on a graphic ratings scale from 1 (Not at All) to 4 (Very Much). The measure was scored by taking the average of item responses. Marteau and Bekker (1992) reported a Cronbach's alpha of .82 for their abbreviated 6-item scale. Higher scores on the scale indicated higher levels of trait anxiety. An example item is "I feel tense". See Appendix F for scale items.

Generalized self-efficacy. I measured generalized self-efficacy using the 8-item New Generalized Self-Efficacy (NGSE) scale (Chen, Gully, & Eden, 2001). Responses were rated on a graphic ratings scale from 1 (Strongly Disagree) to 5 (Strongly Agree). The scale was scored by taking an average of the item responses. Higher scores indicated higher generalized self-efficacy. Chen et al. found the NGSE to have an internal consistency reliability of .86. An

example generalized self-efficacy item is: "I will be able to achieve most of the goals that I have set for myself." See Appendix G for scale items.

Measures to Test Alternative Explanations

Effort. I administered a 20-item measure of subjective effort in the final survey to collect additional data for future research projects but used a 5-item revised version (i.e., Items 1, 9, 11, 12, and 16) between each trial for my analyses. This measure was created and validated in a pilot study (see description below). The Cronbach's alpha for the revised 5-item scale was .87. Participants were asked to rate all items using a graphic rating scale, ranging from 1 (Not at All/None) to 5 (A Lot/Very). I calculated the score for the measure by averaging the item scores. Higher scores on the scale indicated higher subjective perceptions of effort on the current task. An example item is "How much effort are you exerting on the current task". See Appendix H for both the full 20-item scale and the 5-item revised scale items.

I created and validated the above measure of subjective effort in a pilot study. Specifically, I wrote and examined the loadings of 20 items that reflected subjective effort (e.g., "How much effort are you exerting on the current task"; see Appendix H). I collected data from 260 participants. There was complete data for 237 participants due to skipped questions. I conducted an exploratory factor analysis on the 20 items using a Maximum Likelihood extraction method with an oblique rotation (i.e., Direct Oblimin, Delta = 0). I found evidence of three factors meeting the Kaiser criterion with Eigenvalues greater than one. Then, I examined one-, two-, and three-factor solutions. For the one-factor solution, all but one item loaded on the single factor. For the two-factor solution, the fourteen positively keyed items loaded on factor one whereas five of the six reverse-coded items loaded onto the second factor. This indicated that the construct of "not exerting effort" might be distinct from "exerting effort". This non-congruence might be investigated in future research projects. For the three-factor solution, only

one item loaded onto the third factor. That one item also cross-loaded onto another factor (i.e., factor loadings were within .3 of each other).

This indicated that the two-factor solution is the most appropriate solution for the items. However, the second factor, "not exerting effort", is not the construct of interest for this scale. Therefore, I removed the six reverse-coded items and ran an exploratory factor analysis without specifying a factor solution. Two factors had Eigenvalues greater than one. However, the only item that loaded onto the second factor (Item 1) also cross-loaded also onto the first factor. Thus I reran the exploratory factor analysis for a one-factor solution, retaining Item 1. All items loaded on the single factor. Cronbach's alpha was .94 for the thirteen items. Because of the large number of surveys in this thesis, I decided to reduce the item number even further. Because all items had high loadings (i.e., greater than .6), I chose the five items (i.e., Items 1, 9, 11, 12, and 16) I believed best reflected the construct of subjective effort. The Cronbach's alpha for the revised five-item scale was .87.

Cognitive ability. Cognitive ability was measured using the Shipley Institute of Living Scale (SILS) (Shipley, 1940). The SILS is a 60-item scale. A 40-item multiple choice vocabulary section asks the participant to choose one of four responses that has the most similar meaning to a target word. In a 20-item fill-in-the-blank abstraction section, participants must complete a pattern with the correct response. To score the scale, the number of correct answers from each section were summed, with the abstraction correct responses multiplied by two. Higher scores on the test indicated greater cognitive ability. An example vocabulary item is: "TALK" with choices "draw, speak, eat, sleep." An example abstraction item is: "1, 2, 3, 4, 5, ___." Bowers and Pantle (1998) found the SILS to provide similar IQ estimates to other measures of cognitive ability with correlations ranging from .77 to .83. See Appendix I for scale items.

Insufficient effort responding. To account for insufficient effort responding, I inserted the items "Please press 3 if you are reading these instructions", "I have never used a computer", "I work twenty-eight hours in a typical work day", and "Please press 1 if you are reading these instructions" between various measures in both the initial and final survey. "I have never used a computer" and "I work twenty-eight hours in a typical work day" were rated on a 7-point graphic rating scale from (1) Strongly disagree to (7) Strongly agree obtained from Huang and colleagues' (2015) list of infrequency items used to detect careless responding.

Procedure

Participants completed multiple self-report surveys as well as eight trials of a class scheduling task. First, participants completed an informed consent process (see Appendix J). For those choosing to participate, there was an initial survey at the beginning of the study. The initial survey consisted of measures of cognitive ability, optimism, core self-evaluations, trait anxiety, and generalized self-efficacy. I inserted the insufficient effort responding item "Please press 3 if you are reading these instructions" between the core self-evaluations and trait anxiety measures.

Following the initial survey, the participants received instructions for the task. The researcher played an audio recording of the instructions for the task as the participants followed along by reading the instructions on the computer screen. All participants were asked not to move ahead in the instructions on their computer before the audio recording tells them to. All participants completed the instructions at the same time.

Following the instructions, participants reported their self-efficacy for the first trial. After the first trial, participants completed the revised 5-item subjective effort measure referring to the previous trial as well as the self-efficacy measure referring to their upcoming trial. Then, they performed the second trial of the task. Before each of the remaining trials, participants

completed the revised subjective effort measure as well as the self-efficacy measure considering the subsequent trial.

After the final trial, there was an additional survey containing measures of subjective effort (20-item version), the Big Five personality factors, trust in others (facet of agreeableness), modesty (facet of agreeableness), cautiousness (facet of conscientiousness), and demographic information. The insufficient effort responding item (i.e., "Please press 3 if you are reading these instructions") was inserted after the personality measure and before the demographic information. The Big Five personality items were administered in the final survey because personality is shown to be resistant to fatigue. Thus, in both the initial and final survey, I measured dispositional variables in addition to the subjective effort scale in the final survey.

Following the final survey, participants were debriefed and given credit for their participation. See Appendix K for the debriefing form.

Study 2: Comparing SPDs with GPDs and PPDs

Study 2 focused on testing my fourth hypothesis, i.e., that SPDs account for incremental validity in performance beyond that accounted for by GPDs and PPDs.

Participants

Participants in this study were students in an introductory psychology course at a mid-sized Midwestern university. Participation in this study counted toward fulfillment of a requirement in the introductory psychology course.

Experimental Conditions

Participants were randomly assigned to one of four conditions to examine potential treatment and order effects. In Condition 1, participants reported their self-efficacy prior to every trial. In Condition 2, participants set goals prior to each trial. Conditions 1 and 2 allowed me to compare the effects of SPDs, GPDs, and PPDs. In Condition 3, participants set goals for

themselves prior to each of the first four trials but reported their self-efficacy prior to each of the final four trials. In Condition 4, participants reported their self-efficacy prior to each of the first four trials but set goals prior to each of the final four trials. Conditions 3 and 4 are control conditions to enable within subject comparisons of GPDs versus SPDs, controlling for timing of self-efficacy versus goal assessments.

Class Scheduling Task

Participants completed eight trials of the Class Scheduling Task. The Class Scheduling Task (described in Study 1) involves participants navigating a computer interface to create schedules for fictional students with various course preferences. Participants earned 20 points for each completed schedule and were deducted 5 points for each error they make. When errors were corrected, the penalty was removed. Please refer to the description of the Class Scheduling Task in Study 1 for more information.

Measures

Task-specific self-efficacy. I measured task-specific self-efficacy using an 11-item measure of participants' confidence at differing performance levels. The task-specific self-efficacy scale is described in Study 1. See Appendix B for scale items.

Goal Level. I measured goal level using a 1-item measure, asking participants to state the performance level they intend to achieve in the next trial. The participant stated his or her desired cumulative point total for the upcoming trial. Higher scores on the scale indicate higher goal level. The single item was "How many points do you want to earn on the upcoming trial?"

Performance. Performance was operationally defined as the participant's cumulative point total from a given trial of the Class Scheduling Task. Participants earned 20 points for every completed schedule and lost 5 points for each error violation. Points deducted for error violations were replaced for each error the participant fixes.

Self-efficacy – Performance Discrepancy (SPD). I calculated SPDs by taking the absolute value of the difference between the participant's self-efficacy and their actual level of performance for each trial, resulting in eight SPDs.

Goal – Performance Discrepancy (GPD). I calculated the GPDs by taking the absolute value of the difference between the participant's goal level and his/her actual level of performance on each trial, resulting in eight GPDs.

Past-performance – Performance Discrepancy (PPD). I calculated PPDs by taking the absolute value of the difference between performance on a trial and performance on the prior trial. The first PPD will need a prior performance level, so PPD_2 was computed as the difference between performance on Trial 2 and Trial 1. Over the course of eight trials, there were a total of seven PPDs.

Demographics. I measured demographics in a survey asking participants about sex, age, year, major, GPA, ethnicity, and handedness (see Appendix A). I asked participants about their handedness because the class scheduling task involved near exclusive use of the right hand. If people are left-handed, they might perform poorly on the task simply by using their non-dominant hand.

Measures to Test Alternative Explanations

I included the following measures to test alternative explanations for my results.

Personality. I measured personality using the 50-item measure of the Big Five personality factors and 28 additional items from Costa and McCrae's (1992) NEO-PI-R IPIP measure for a total of 78 personality items (International Personality Item Pool, 2013). This measure is described in Study 1. See Appendix C for scale items.

Optimism. I measured dispositional optimism using the 10-item Life Orientation Test-Revised (LOT-R) optimism measure developed by Scheier, Carver, and Bridges (1994). This measure is described in Study 1. Appendix D for scale items.

Core Self-Evaluations. I measured core self-evaluations using the 12-item Core Self-Evaluations Scale (CSES) measure developed by Judge, Erez, Bono, and Thoresen (2003). This measure is described in Study 1. See Appendix E for scale items.

Trait anxiety. I measured trait anxiety using the 6-item abbreviated version of Spielberger's (1983) State/Trait Anxiety Inventory (STAI) scale developed by Marteau and Bekker (1992). This measure is described in Study 1. See Appendix F for scale items.

Generalized self-efficacy. I measured generalized self-efficacy using the 8-item New Generalized Self-Efficacy (NGSE) scale (Chen, Gully, & Eden, 2001). This measure is described in Study 1. See Appendix G for scale items.

Effort. I administered a 20-item measure of subjective effort to collect additional data for future research projects but used a 5-item revised version between each trial for my analyses. The validation steps and information about the measure is described in Study 1. See Appendix H for scale items.

Cognitive ability. I measured cognitive ability using the Shipley Institute of Living Scale (SILS) (Shipley, 1940). This measure is described in Study 1. An example abstraction item is: "1, 2, 3, 4, 5, ___.". See Appendix I for scale items.

Perceived control. I measured perceived control by using a single-item measure I generated to examine possible effects of instrumentality perceptions on the skill acquisition process. The item was "To what extent did you believe you had the ability/control/resources to improve your performance to match your expected/goal cumulative score?".

Insufficient effort responding. To account for insufficient effort responding, I included four items: “Please press 3 if you are reading these instructions”, “I have never used a computer”, “I work twenty-eight hours in a typical work day”, and “Please press 1 if you are reading these instructions” between various measures in both the initial and final survey.

Procedure

Participants completed multiple self-report surveys as well as eight trials of a class scheduling task. First, participants completed an informed consent process (see Appendix J). For those choosing to participate, there was an initial survey at the beginning of the study. The initial survey consisted of measures of cognitive ability, optimism, core self-evaluations, trait anxiety, and generalized self-efficacy. I inserted the insufficient effort responding item “Please press 3 if you are reading these instructions” between the core self-evaluations and trait anxiety measures.

Following the initial survey, the participants learned the instructions for the task. The researcher played an audio recording of the instructions for the task as the participants followed along by reading the instructions on the computer screen. All participants were asked not to move ahead in the instructions on their computer before the audio recording tells them to. All participants completed the instructions at the same time. The class scheduling task is described in the Method section of Study 1.

Following the instructions, participants either reported their self-efficacy or goal level for the upcoming trial. After the first trial, participants completed the revised 5-item subjective effort measure referring to the previous trial as well as the self-efficacy or goal-level measure referring to their upcoming trial. Then, they performed the second trial of the task. Before each of the remaining trials, participants completed the revised subjective effort measure as well as the self-efficacy or goal-level measure considering the subsequent trial.

After the final trial, there was an additional survey for participants to complete containing measures of subjective effort (20-item version), the Big Five personality factors, trust in others (facet of agreeableness), modesty (facet of agreeableness), cautiousness (facet of conscientiousness), and demographic information. The insufficient effort responding item (i.e., "Please press 3 if you are reading these instructions") was inserted after the personality measure and before the demographic information. The Big Five personality items were administered in the final survey because personality is shown to be resistant to fatigue.

Following the final survey, participants were debriefed and given credit for their participation. See Appendix K for the debriefing form.

Results

Data Cleaning

Study 1. Of the 176 participants, 27 participants had partial survey response data removed due to insufficient effort responding. There were two surveys, one before the task and one after. In each survey, I included an IER item halfway through the survey and another at the end. To retain as much data as possible, if a participant were to trigger an IER item, I only removed their responses that followed the most recent IER item. For example, if a participant triggered the IER item at the end of the survey but not the middle, I only removed their responses between the middle IER item and the end IER item. I did this to retain as much data as possible. To detect insufficient effort responding, I checked whether participants failed the insufficient effort responding items. I planned to remove participants from analyses if they were either under 18 or reported not being fluent in English at all. However, due to a programming error, I was unable to collect demographic information for Study 1. Consequently, I was unable to clean Study 1 based on age and/or English fluency.

Study 2. Of the 321 participants, I excluded 2 for reporting their age as 17, as being 18 was a prerequisite for participating in the study. One participant was removed for reporting him/herself as a graduate student. I also removed all self-reported data from 1 participant because s/he reported he/she was not fluent in English at all. Partial data was removed from the survey responses due to IER from 50 participants. All impossible values were deleted and became missing values. I checked for outliers by searching for scores higher or lower than 4 standard deviations away from the mean score. Any scores farther than 4 standard deviations away from the mean were treated as missing data in subsequent analyses.

Psychometric Properties of Measures.

I tested internal consistency reliabilities for all measures used in Study 1 and Study 2. For every scale, there was not a single item that, if removed, would improve the internal consistency in either sample. For this reason, I chose not to remove any items when conducting subsequent analyses. The internal consistency for the Scheier, Carver, and Bridges (1994) optimism scale was .70 in Study 1 and .71 in Study 2. Judge, Erez, Bono, and Thorensen's (2003) Core Self-Evaluations scale had internal consistency reliability estimates of .88 in Study 1 and .84 in Study 2. The trait anxiety scale (Spielberger, 1983) had an internal consistency reliability of .83 in Study 1 and .81 in Study 2. The New Generalized Self-Efficacy (NGSE) scale (Chen, Gully, & Eden, 2001) had internal consistency reliability estimates of .92 in Study 1 and .90 in Study 2. The 5-item effort scale developed for this study was administered eight times, demonstrating minimum internal consistency reliability estimates across all eight administrations of .87 in Study 1 and .89 in Study 2. The personality measures from the NEO-PI-R IPIP (Costa & McCrae, 1992) showed similar internal consistency estimates in both Study 1 and Study 2, respectively (Neuroticism: .85, .83; Extraversion: .86, .84; Openness: .78, .79;

Agreeableness: .72, .71; Conscientiousness: .77, .76; Trust: .80, .80; Modesty: .71, .74; Cautiousness: .75, .77).

Descriptive Statistics

Due to a programming error, demographic information was not collected for Study 1. However, samples in Study 1 and 2 should be demographically similar because they both consisted of Introduction to Psychology students completing the study for course credit. The mean age in Study 2 was 19.03 years with a standard deviation of 2.39 years. In Study 2, 64.5% of participants were female, 32.8% were male, 1.7% identified as non-binary/third gender, and 1% preferred not to say. The most common self-reported ethnicity in Study 2 was White/Caucasian (75.7%), followed by Black/African American (15.2%), Asian/Pacific (3.7%), Hispanic (2.7%), Native American (.7%), and additional unlisted ethnicities (2%). Most participants were freshmen (77.4%), followed by sophomores (12.5%), juniors (8.1%), seniors (2%). Of the participants, 87.2% were right-handed, 6.4% left-handed, and 6.4% were ambidextrous. When asked how fluent in English they were, 93.1% of participants reported being very fluent, 2.8% reported being somewhat fluent, and .3% reported being not fluent at all. Data from those who reported being not fluent in English was removed.

To visualize the progression of my key variables across trials, I obtained the means, standard deviations, and correlations in both Study 1 and 2 (see Tables 1, 2, and 3).

Table 1

Progression of Key Study Variables Over Time (Study 1)

	Trial															
	1		2		3		4		5		6		7		8	
Variables	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Self-efficacy	52	88	101	108	125	101	157	98	161	100	175	101	183	107	193	103
Performance	53	47	125	66	164	73	187	84	209	95	217	96	227	107	228	112
SPD	0	104	-29	112	-40	103	-33	109	-51	114	-46	98	-45	96	-39	99
|SPD|	71	76	95	65	89	65	91	68	100	75	85	67	85	64	81	69
PPD			72	46	38	40	23	40	21	42	9	42	10	50	1	45

Note. Values rounded to the nearest whole number

Table 2

Progression of Key Study Variables Over Time (Study 2)

	Trial															
	1		2		3		4		5		6		7		8	
Variables	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Self-efficacy	43	81	85	101	111	98	127	102	137	103	161	103	161	103	159	102
Performance	50	45	110	64	148	70	171	82	192	84	195	85	205	90	204	94
Goal level	60	55	61	43	90	67	110	76	115	75	137	83	145	89	149	92
SPD	-5	95	-27	104	-38	105	-49	104	-60	103	-38	102	-48	102	-50	96
|SPD|	63	72	83	68	88	67	89	72	94	73	83	71	85	73	82	71
GPD	7	62	-46	61	-55	49	-50	64	-67	65	-54	65	-50	68	-50	68
PPD			61	41	37	41	22	42	22	44	4	50	8	45	0	48

Note. Values rounded to the nearest whole number

Table 3

Correlations Between Key Study Variables Over Time

	Trial							
Variables	1	2	3	4	5	6	7	8
Study 1								
Self-efficacy ~~ Performance	-.11	.23**	.34**	.29**	.31**	.51**	.60**	.58**
Study 2								
Self-efficacy ~~ Performance	-.06	.26**	.27**	.39**	.42**	.42**	.47**	.53**
Goal level ~~ Performance	.25*	.41**	.73**	.66**	.64**	.70**	69**	.72**

Note. * = $p < .05$, ** = $p < .01$; correlations between Self-efficacy and Goal level were unavailable because there were no individuals that reported both variables on a given trial.

Hypothesis Testing

Because I examined participants' change in outcomes over time, I used multi-level modeling to test my hypotheses. I followed the analysis approach defined by Bliese and Ployhart (2002) and describe the steps I followed below. My only Level 1 predictor was trial. All other predictors were Level 2 (e.g., initial SPD, optimism, agreeableness). The outcome was $|SPD|_{2\text{-}8}$. When testing these hypotheses for Study 2, I used all three conditions that produced SPDs (a) throughout all 8 trials, (b) in the first 4 trials, or (c) in the last 4 trials.

Hypothesis 1. In Hypothesis 1, I proposed that trial would have a negative effect on absolute SPD, i.e., that the absolute size of the SPD would decrease across trials. First, I

estimated intraclass correlation coefficients for both Study 1 and Study 2 to determine what amount of variance in outcomes was at the between person level. Both ICCs were greater than .1 (Study 1 = .66, Study 2 = .65), indicating that a substantial amount of total variance in |SPD| was at the between person level. Next, I used multilevel modeling to examine whether there was variance in intercepts and/or slopes across individuals. To do so, I tested the simplest model, which allowed only the intercepts to vary:

$$Y_{ij} = ß_{0j} + ß_{1j}x_{ij} + r_{ij}$$

$$ß_{0j} = \gamma_{00} + \mu_{0j}$$

$$ß_{1j} = \gamma_{10}$$

where Y_{ij} represents SPD on trial *i* for participant *j*, and x_{ij} represents trial *i* for participant *j*. The intercept ($ß_{0j}$) and the slope ($ß_{1j}$) are represented by the common intercept (γ_{00}) and slope (γ_{10}), respectively, across individuals. Finally, μ_{0j} represents variance in the intercept.

Then, I created models for Study 1 and Study 2 that allowed both intercepts and slopes to vary:

$$Y_{ij} = ß_{0j} + ß_{1j}x_{ij} + r_{ij}$$

$$ß_{0j} = \gamma_{00} + \mu_{0j}$$

$$ß_{1j} = \gamma_{10} + \mu_{1j}$$

where I've now added μ_{1j}, which represents variance in the slope.

For Study 1, the deviance score for the random intercept model (deviance = -6446.99) was significantly different from the deviance score for the random intercept/slope model (deviance = -6418.12), X^2_{diff} (2) = 57.73, $p < .01$. For Study 2, the deviance score for the random intercept model (deviance = -7356.79) was significantly different from the deviance score for the random intercept/slope model (deviance = -7332.46), X^2_{diff} (2) = 48.66, $p < .01$. Thus, for both

studies, the model allowing intercepts and slopes to vary fit significantly better than the model allowing only intercepts to vary.

Subsequently, I conducted tests of autocorrelation to determine whether responses that were closer in time were more strongly related to each other than responses that were farther apart. I found evidence of autocorrelation, and after controlling for it, the model for Study 1 and Study 2 fit significantly better. For Study 1, the deviance score for the random intercept/slope model (deviance = -6418.12) was significantly different from the deviance score for the autocorrelation model (deviance = -6384.59), X^2_{diff} (21) = 67.04, $p < .01$. For Study 2, the deviance score for the random intercept/slope model (deviance = -7332.46) was significantly different from the deviance score for the autocorrelation model (deviance = -7313.67), X^2_{diff} (21) = 37.58, $p < .05$. Thus, for both studies, the model accounting for autocorrelation fit significantly better than the random intercepts/slopes models.

Finally, I tested for heteroscedasticity, i.e., whether residuals for different cases have different variances, using a Levene test of homogeneity. I found evidence of heteroscedasticity in the |SPD| outcome in both studies, Study 1: F (175, 991) = 1.23, $p < .05$; Study 2: F (247, 1086) = 1.40, $p < .01$. Then, I controlled for heteroscedasticity for both studies and found a significant improvement. For Study 1, the deviance score for the autocorrelation model (deviance = -6384.59) was significantly different from the deviance score for the heteroscedasticity model (deviance = -6378.11), X^2_{diff} (6) = 12.98, $p < .05$. For Study 2, the deviance score for the autocorrelation model (deviance =-7313.67) was significantly different from the deviance score for the heteroscedasticity model (deviance = -7301.69), X^2_{diff} (6) = 23.97, $p < .01$. Thus, for both studies, the model accounting for heteroscedasticity fit significantly better than the autocorrelation models.

Using the best fitting models for each study, I tested whether $ß_{1j}$ was negative and statistically significant. I obtained support for Hypothesis 1 because time significantly predicted |SPD| in Study 1 ($ß_{1j}$ = -2.14, $p < .05$) and Study 2 ($ß_{1j}$ = -1.96, $p < .05$). Results indicated that |SPD| decreased over time in both Study 1 and Study 2 (see Table 4, Figure 2, and Figure 3).

Table 4

Main Effects of Time on |SPD_{2-8}|

Fixed effects	*b*	*SE*	*df*	*t*	*p*
Study 1					
Time	-2.14	1.03	990	-2.08	.037
Study 2					
Time	-1.96	.98	1085	-2.00	.045

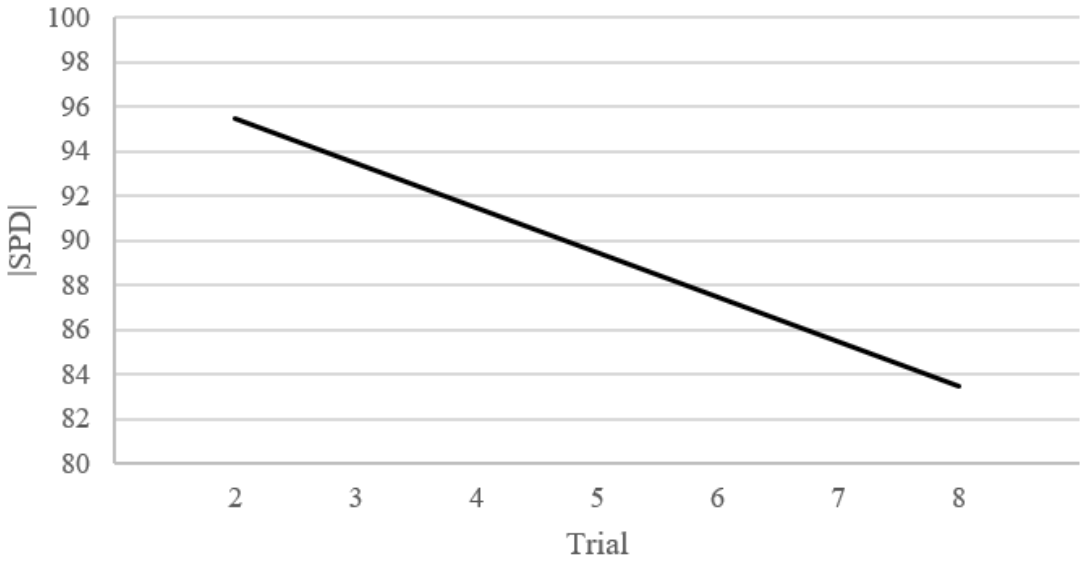

Figure 2. The decrease of |SPD| over time (Study 1)

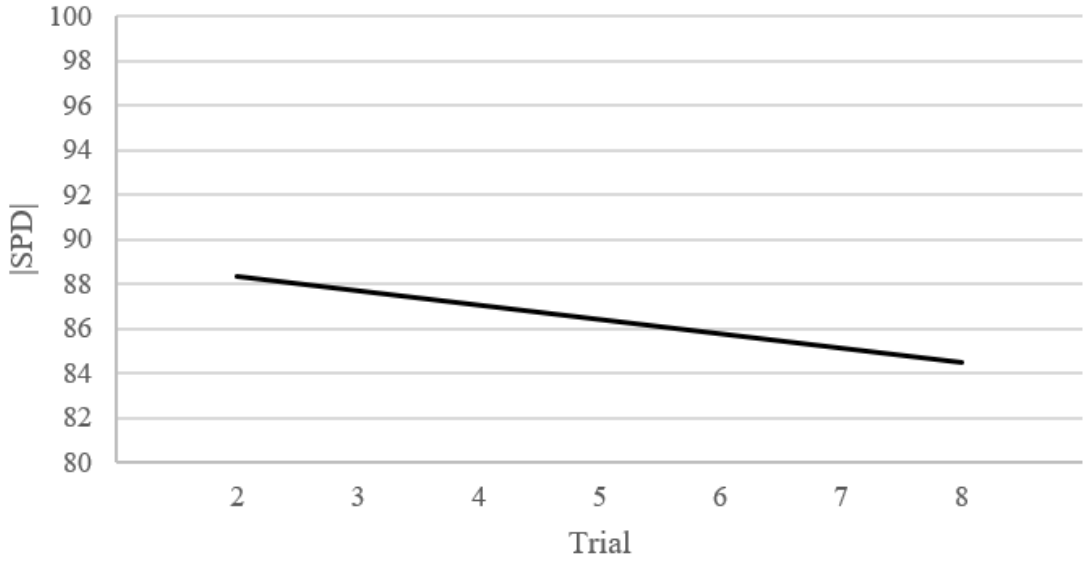

Figure 3. The decrease of |SPD| over time (Study 2)

Hypothesis 2. I used the same approach to test Hypothesis 2, which posited that trial and the sign of SPD_1 (initial SPD) interact in their effect on absolute SPD (in subsequent trials) such that those initially overestimating their ability level would reduce their |SPD|s more quickly over time. The sign of the initial SPD was dummy coded such that a negative SPD_1 was coded as 0 and a positive SPD_1 was coded as 1. To test the interaction of trial (Level 1) with sign of SPD_1 (Level 2), I used the following equations:

$$Y_{ij} = ß_{0j} + ß_{1j}x_{ij} + r_{ij}$$

$$ß_{0j} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j}$$

$$ß_{1j} = \gamma_{10} + \gamma_{11}(z_j) + \mu_{1j}$$

where Y_{ij} is SPD for trial *i* for individual *j*, x_{ij} is trial *i* for individual *j*, and z_j is the sign of SPD_1 for participant *j*. The full equation was:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j} + \gamma_{10}(x_{ij}) + \gamma_{11}(z_j)(x_{ij}) + \mu_{1j}(x_{ij}) + r_{ij}$$

Thus, the coefficient for the Trial X sign of SPD_1 interaction term is γ_{11}.

I used the models described in Hypothesis 1, adding the Sign of SPD main effect and an interaction term (Trial X SignSPD_1) to the fitted models. The interaction term, γ_{11}, was non-significant in Study 1 (γ_{11} = -1.90, p = .39) and Study 2 (γ_{11} = -.24, p = .92; see Table 5). These results did not provide support for Hypothesis 2. These findings indicated that |SPD|s did not decrease at a different rate depending on the sign of the initial SPD.

Table 5

Interactive Effects of Time X SignSPD_1 Predicting |SPD_{2-8}|

Fixed effects	*b*	*SE*	*df*	*t*	*p*
Study 1					
Time	-2.13	1.33	906	-1.60	.109
signSPD_1	-.12	12.70	159	-.01	.993
Time X signSPD_1	-1.90	2.20	906	-.86	.389
Study 2					
Time	-2.20	1.30	827	-1.69	.092
signSPD_1	-8.29	12.70	173	-.65	.515
Time X signSPD_1	-.24	2.28	827	-.10	.917

Hypothesis 3. Hypothesis 3 posited that trial and the size of SPD_1 ($|SPD_1|$) interact in their effect on absolute SPD (in subsequent trials) such that those who initially under-/overestimate their ability level to a larger extent would reduce their |SPD|s more quickly over time. The size of SPD_1 was calculated as the absolute value of the first SPD. To test the interaction of trial (Level 1) with size of SPD_1 (Level 2), I used the following equations:

$$Y_{ij} = ß_{0j} + ß_{1j}x_{ij} + r_{ij}$$

$$ß_{0j} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j}$$

$$ß_{1j} = \gamma_{10} + \gamma_{11}(z_j) + \mu_{1j}$$

where Y_{ij} is SPD for trial *i* for individual *j*, x_{ij} is trial *i* for individual *j*, and z_j is the size of SPD_1 for participant *j*. The full equation is:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j} + \gamma_{10}(x_{ij}) + \gamma_{11}(z_j)(x_{ij}) + \mu_{1j}(x_{ij}) + r_{ij}$$

Thus, the coefficient for the Trial X $|SPD_1|$ interaction term is γ_{11}. I used the models described in Hypothesis 1, adding the $|SPD_1|$ main effect and an interaction term (Trial X $|SPD_1|$) to the fitted models. The interaction term, γ_{11}, was not statistically significant in Study 1 (γ_{11} = -.02, p = .09) but was significant in Study 2 (γ_{11} = -.03, $p < .05$; see Table 6). These results provided partial support for Hypothesis 3, that trial number and $|SPD_1|$ interact when predicting subsequent |SPD|. These findings indicated that |SPD|s increased at a slower rate for those with larger $|SPD_1|$s for Study 2, but not in Study 1. Interestingly, the main effect of $|SPD_1|$ when predicting subsequent |SPD|s was significant in both Study 1 (γ_{01} = .26, $p < .01$) and Study 2 (γ_{01} = .26, $p < .01$).

Table 6

Interactive Effects of Time X |SPD$_1$| Predicting |SPD$_{2\text{-}8}$|

Fixed effects	*b*	*SE*	*df*	*t*	*p*
Study 1					
Time	-.76	1.45	954	-.52	.603
|SPD$_1$|	.26	.08	167	3.20	.002
Time X |SPD$_1$|	-.02	.01	954	-1.68	.094
Study 2					
Time	.20	1.40	939	10.33	.885
|SPD$_1$|	.26	.08	197	3.26	.001
Time X |SPD$_1$|	-.03	.02	939	-2.30	.022

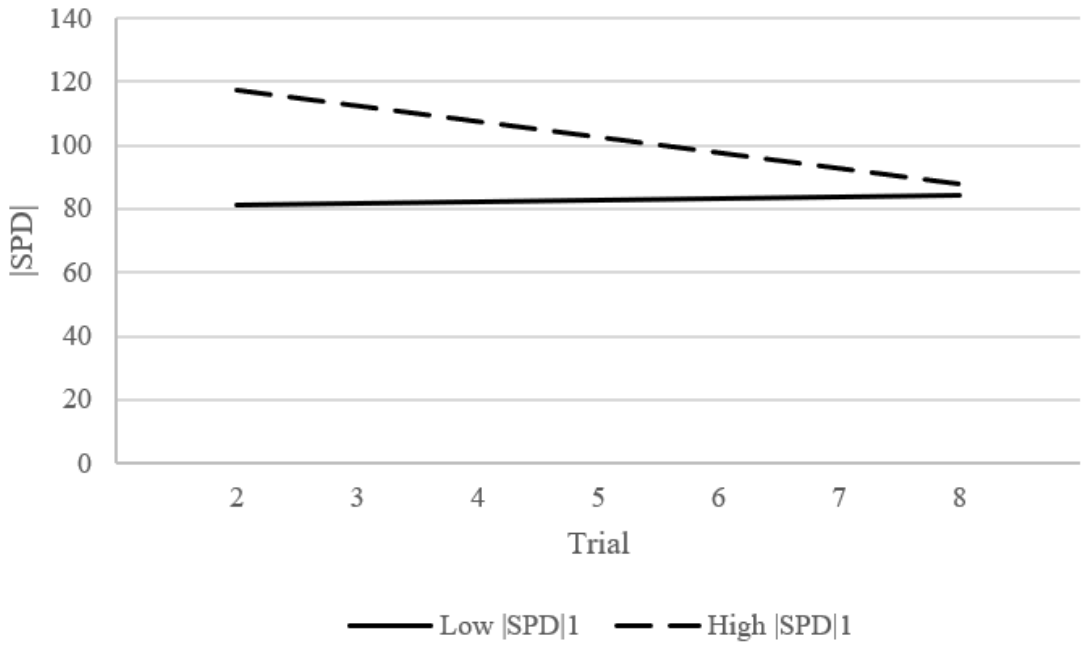

Figure 4. The interaction between |SPD|$_1$ and Trial (Study 2)

Research Question. In the research question, I examined whether personality factors interacted with trial when predicting |SPD|s such that higher levels of various personality characteristics would lead to a quicker reduction of |SPD|s over time. Because I tested eight different personality variables as moderators (i.e., agreeableness, trust, modesty, cautiousness, optimism, core self-evaluations, trait anxiety, and generalized self-efficacy), I used multi-level modeling to test each of the eight personality variables individually. To test the interaction of trial (Level 1), size of SPD_1 (Level 2), and personality factors (Level 2), I used the following equations:

$$Y_{ij} = ß_{0j} + ß_{1j}x_{ij} + r_{ij}$$

$$ß_{0j} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j}$$

$$ß_{1j} = \gamma_{10} + \gamma_{11}(z_j) + \mu_{1j}$$

where Y_{ij} is SPD for trial *i* for individual *j*, x_{ij} is trial *i* for individual *j*, and z_j is the personality variable for participant *j*. The full equation is:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j} + \gamma_{10}(x_{ij}) + \gamma_{11}(z_j)(x_{ij}) + \mu_{1j}(x_{ij}) + r_{ij}$$

Thus, the coefficient for the Trial X Personality interaction term is γ_{11}. All but one interaction was non-significant, indicating that most of the selected personality variables failed to interact with trial when predicting subsequent |SPD|s (see Table 7).

Table 7

Personality Interactions Predicting $|SPD_{2\text{-}8}|$

	Study 1		Study 2	
Interactions	*b*	*p*	*b*	*p*
Agreeableness X Time	-1.11	.56	.08	.96
Trust X Time	-.61	.70	1.26	.45
Modesty X Time	2.71	.11	-.62	.69
Cautiousness X Time	-.20	.90	.34	.86
Optimism X Time	-1.14	.42	-2.36	.08
Core Self-evaluations X Time	-.77	.61	-.79	.63
Trait anxiety X Time	1.37	.38	.48	.75
Generalized Self-efficacy X Time	-3.32	.03	-1.18	.47

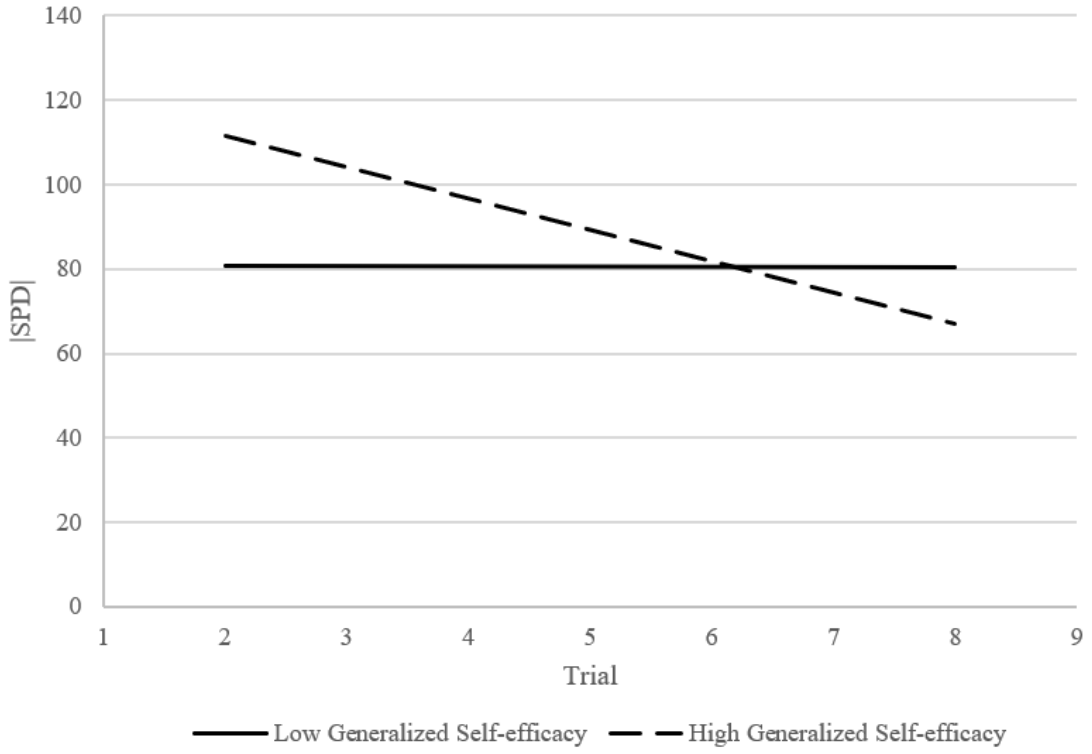

Figure 5. The interaction between Generalized Self-efficacy and Trial (Study 1)

Hypothesis 4. Hypothesis 4 suggested that SPDs would account for incremental variance in performance beyond the variance accounted for by GPDs and PPDs. Due to the experimental design, I was unable to test Hypothesis 4 as it was originally conceived. This is because to test for incremental variance, there must exist cases where both GPDs and SPDs exist within a participant on a trial. Because of concerns about reactivity effects between GPDs and SPDs, I set up the experiment so that only one of the two would be obtained each trial. Therefore, each participant was assigned to one of four conditions: (1) $SPD_{1\text{-}8}$, (2) $GPD_{1\text{-}8}$, (3) $SPD_{1\text{-}4}$ and $GPD_{5\text{-}8}$, and (4) $GPD_{1\text{-}4}$ and $SPD_{5\text{-}8}$.

Instead of testing for incremental variance to detect differences between the discrepancy variables, I tested four predictive models of performance (see Table 8). Each model included Condition to capture differences between SPDs and GPDs.

Table 8

Hypothesis 4 Models

Models
(a) condition (1 v 2) X trial + condition (1 v 2) X trial^2
(b) condition (3 v 4) X trial + condition (3 v 4) X trial^2
(c) condition (1 v 2) X trial X PPD_2 + condition (1 v 2) X trial^2 X PPD_2
(d) condition (3 v 4) X trial X PPD_2 + condition (3 v 4) X trial^2 X PPD_2

The first two models were meant to examine whether the linear or quadratic rate of learning varied due to condition, and the second two models added PPD_2 (i.e., the difference in performance from Trial 1 to 2) to examine whether condition was a significant predictor of performance in the presence of prior performance. Condition 1 was SPDs only, Condition 2 was GPDs only, Condition 3 was SPDs for Trials 1-4 and GPDs for Trials 5-8, and Condition 4 was GPDs for Trials 1-4 and SPDs for Trials 5-8. PPD_2 was defined as the difference between performance levels on trial 1 and 2. Therefore, the outcome for the interactions was performance on Trials 3-8.

Hypothesis 4a. First, I investigated possible differences between Condition 1 and 2, which would indicate whether a difference in performance depended on participants' reported self-efficacy or goal level before each trial. Therefore, to test the interactions of trial X condition (1v2), and trial^2 X condition (1 v 2), I used the following equations:

$$Y_{ij} = ß_{0j} + ß_{1j}x_{ij} + ß_{2j}x_{ij}^2 + r_{ij}$$

$$ß_{0j} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j}$$

$$ß_{1j} = \gamma_{10} + \gamma_{11}(z_j) + \mu_{1j}$$

$$ß_{2j} = \gamma_{20} + \gamma_{21}(z_j) + \mu_{2j}$$

where Y_{ij} is performance for trial *i* for individual *j*, x_{ij} is trial *i* for individual *j*, and z_j is condition (1 or 2) for participant *j*. The full equation is:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j} + \gamma_{10}(x_{ij}) + \gamma_{11}(z_j)(x_{ij}) + \mu_{1j}(x_{ij}) + \gamma_{20}(x_{ij}^2) + \gamma_{21}(z_j)(x_{ij}^2) + \mu_{2j}(x_{ij}^2) + r_{ij}$$

Thus, the coefficient for the trial X condition (1 or 2) interaction term is γ_{11} and the coefficient for the trial^2 X condition (1 or 2) interaction term is γ_{21}. The interaction between trial and condition (1 or 2) was significant ($\gamma_{11} = -21.85$, $p < .05$; see Table 9), indicating that performance increased quicker for those reporting self-efficacy before each trial compared to those reporting goal levels. Also, the interaction between trial^2 and condition was significant ($\gamma_{21} = 1.95$, $p < .05$), indicating that those reporting self-efficacy before each trial increased performance quicker but also saw a dip in performance in the last few trials (see Figure 6). These results suggested that the rate at which someone learns a new task varies depending on whether they report self-efficacy or goal levels before each trial. Interestingly, when the same model was tested predicting $\text{Performance}_{1\text{-}8}$ instead of $\text{Performance}_{3\text{-}8}$, the significant condition interactions became non-significant, indicating that the difference in performance between those in the SPD-only and GPD-only condition was exhibited only in later stages of skill acquisition.

Table 9

Interactive Effects of Time X Condition (1 v 2) and Time^2 X Condition (1 v 2) Predicting Performance$_{3-8}$

Fixed effects	*b*	*SE*	*df*	*t*	*p*
Time	55.36	4.89	1020	11.33	.000
Condition (1 v 2)	40.05	24.87	205	1.61	.109
Time^2	-4.05	.44	1020	-9.17	.000
Time X Condition (1 v 2)	-21.85	9.98	1020	-2.19	.029
Time^2 X Condition (1 v 2)	1.95	.90	1020	2.16	.031

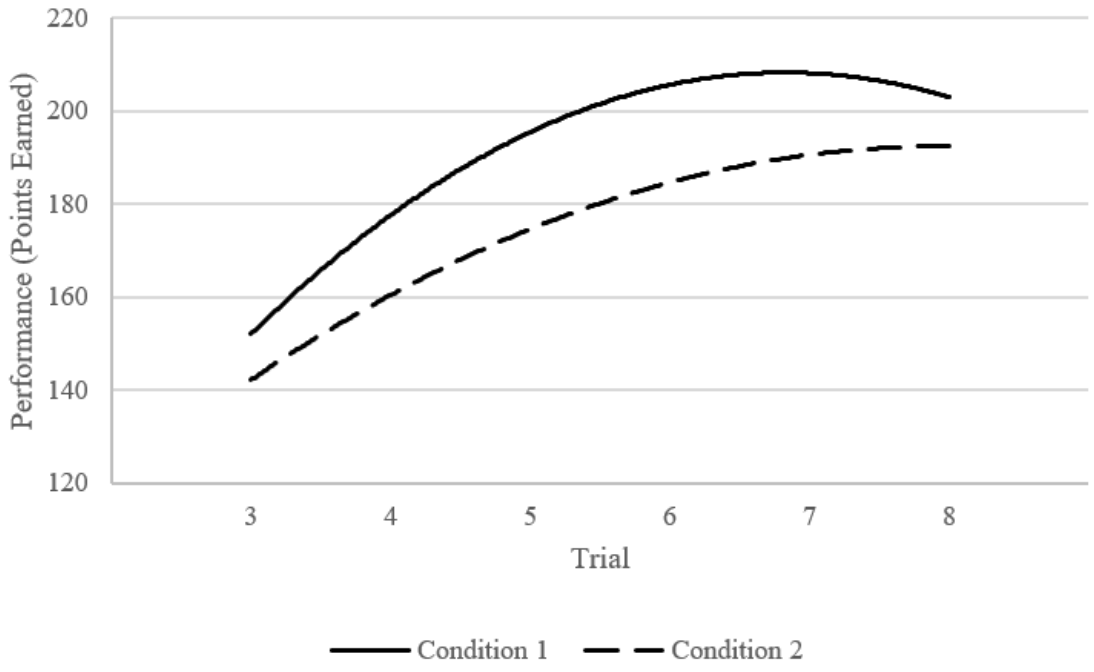

Figure 6. The quadratic interaction of Condition (1 versus 2) and Trial (Study 2)

Hypothesis 4b. Next, I investigated possible differences between Condition 3 and 4, which would indicate whether the order in which participants reported self-efficacy or goal level

before each trial mattered when predicting performance. Therefore, to test the interactions of trial X condition (3 v 4) and trial^2 X condition (3 v 4), I used the following equations:

$$Y_{ij} = ß_{0j} + ß_{1j}x_{ij} + ß_{2j}x_{ij}^2 + r_{ij}$$

$$ß_{0j} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j}$$

$$ß_{1j} = \gamma_{10} + \gamma_{11}(z_j) + \mu_{1j}$$

$$ß_{2j} = \gamma_{20} + \gamma_{21}(z_j) + \mu_{2j}$$

where Y_{ij} is performance for trial *i* for individual *j*, x_{ij} is trial *i* for individual *j*, and z_j is condition (3 or 4) for participant *j*. The full equation is:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(z_j) + \mu_{0j} + \gamma_{10}(x_{ij}) + \gamma_{11}(z_j)(x_{ij}) + \mu_{1j}(x_{ij}) + \gamma_{20}(x_{ij}^2) + \gamma_{21}(z_j)(x_{ij}^2) + \mu_{2j}(x_{ij}^2) + r_{ij}$$

Thus, the coefficient for the Trial X condition (3 or 4) interaction term is γ_{11} and the coefficient for the Trial^2 X condition (3 or 4) interaction term is γ_{21}. Both interactions were non-significant (γ_{11} = -13.02, p = .32; γ_{21} = 1.46, p = .24; see Table 10), indicating no significant difference in performance between those producing SPDs first followed by GPDs vs. those producing GPDs first followed by SPDs.

Table 10

Interactive Effects of Time X Condition (3 v 4) and Time^2 X Condition (3 v 4) Predicting Performance$_{3-8}$

Fixed effects	*b*	*SE*	*df*	*t*	*p*
Time	39.45	9.15	456	4.31	.000
Condition (3 v 4)	33.28	32.30	91	1.03	.306
Time^2	-2.53	.87	456	-2.90	.004
Time X Condition (3 v 4)	-13.02	13.07	456	-.99	.320
Time^2 X Condition (3 v 4)	1.46	1.24	456	1.17	.241

Hypothesis 4c. Next, I included past-performance – current performance discrepancies (PPDs). I examined whether the SPD/GPD motivational effects still existed in the presence of prior performance, the strongest predictor of subsequent performance. So, I decided to test the effect of PPD_2, the difference in performance from trial 1 to 2, to determine whether condition (1 or 2) accounted for unique variance in performance beyond that accounted for by the initial change in performance. Therefore, to test the interactions of trial X condition (1 or 2) and trial^2 X condition (1 or 2), I used the following equations:

$$Y_{ij} = ß_{0j} + ß_{1j}x_{ij} + ß_{2j}x_{ij}^2 + r_{ij}$$

$$ß_{0j} = \gamma_{00} + \gamma_{01}(z_j) + \gamma_{02}(m_j) + \gamma_{03}(z_j)(m_j) + \mu_{0j}$$

$$ß_{1j} = \gamma_{10} + \gamma_{11}(z_j) + \gamma_{12}(m_j) + \gamma_{13}(z_j)(m_j) + \mu_{1j}$$

$$ß_{2j} = \gamma_{20} + \gamma_{21}(z_j) + \gamma_{22}(m_j) + \gamma_{23}(z_j)(m_j) + \mu_{2j}$$

Where Y_{ij} is performance for trial *i* for individual *j*, x_{ij} is trial *i* for individual *j*, z_j is Condition (1 or 2) for participant *j*, m_j is PPD_2 for participant *j*. The full equation is:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(z_j) + \gamma_{02}(m_j) + \gamma_{03}(z_j)(m_j) + \mu_{0j} + \gamma_{10}(x_{ij}) + \gamma_{11}(z_j)(x_{ij}) + \gamma_{12}(m_j)(x_{ij}) + \gamma_{13}(z_j)(m_j)(x_{ij})$$
$$+ \mu_{1j}(x_{ij}) + \gamma_{20}(x_{ij}^2) + \gamma_{21}(z_j)(x_{ij}^2) + \gamma_{22}(m_j)(x_{ij}^2) + \gamma_{23}(z_j)(m_j)(x_{ij}^2) + \mu_{2j}(x_{ij}^2) + r_{ij}$$

Thus, the coefficient for the trial X condition (1 or 2) interaction was γ_{11} and the coefficient for the trial^2 X condition (1 or 2) is γ_{21}. Neither interaction was statistically significant (γ_{11} = -20.09, p = .19; γ_{21} = 1.81, p = .18; see Table 11), indicating that condition did not account for incremental variance beyond that accounted for by the initial change in performance, PPD_2. This provides further evidence that past performance is the best predictor of subsequent performance, and if prior performance is known, other predictors such as GPDs or SPDs do not additional information.

Table 11

Interactive Effects of Time X Condition (1 v 2) and Time^2 X Condition (1 v 2) Predicting Performance$_{3-8}$

Fixed effects	*b*	*SE*	*df*	*t*	*p*
Time	51.34	9.13	1010	5.62	.000
Condition (1 v 2)	57.34	40.18	201	1.43	.155
PPD_2	.91	.28	201	3.22	.002
$Time^2$	-3.64	.83	1010	-4.40	.000
Time X Condition (1 v 2)	-20.09	15.43	1010	-1.30	.193
Time X PPD_2	.05	.12	1010	.46	.648
Condition (1 v 2) X PPD_2	-.17	.59	201	-.30	.766
$Time^2$ X Condition (1 v 2)	1.81	1.35	1010	1.34	.181
PPD_2 X $Time^2$	-.01	.01	1010	-.53	.594
Time X Condition (1 v 2) X PPD_2	.01	.22	1010	.04	.970
$Time^2$ X Condition (1 v 2) X PPD_2	-.00	.02	1010	-.06	.956

Hypothesis 4d. Similarly, I examined the effects of initial past-performance – current performance discrepancies (i.e., PPD_2) in conjunction with condition 3 or 4 when predicting performance. Therefore, to test the interactions of trial X condition (3 or 4) and trial^2 X condition (3 or 4), I used the following equations:

$$Y_{ij} = ß_{0j} + ß_{1j}x_{ij} + ß_{2j}x_{ij}^2 + r_{ij}$$

$$ß_{0j} = \gamma_{00} + \gamma_{01}(z_j) + \gamma_{02}(m_j) + \gamma_{03}(z_j)(m_j) + \mu_{0j}$$

$$ß_{1j} = \gamma_{10} + \gamma_{11}(z_j) + \gamma_{12}(m_j) + \gamma_{13}(z_j)(m_j) + \mu_{1j}$$

$$ß_{2j} = \gamma_{20} + \gamma_{21}(z_j) + \gamma_{22}(m_j) + \gamma_{23}(z_j)(m_j) + \mu_{2j}$$

where Y_{ij} is performance for trial *i* for individual *j*, x_{ij} is trial *i* for individual *j*, z_j is Condition (3 or 4) for participant *j*, m_j is PPD_2 for participant *j*. The full equation is:

$$Y_{ij} = \gamma_{00} + \gamma_{01}(z_j) + \gamma_{02}(m_j) + \gamma_{03}(z_j)(m_j) + \mu_{0j} + \gamma_{10}(x_{ij}) + \gamma_{11}(z_j)(x_{ij}) + \gamma_{12}(m_j)(x_{ij}) + \gamma_{13}(z_j)(m_j)(x_{ij})$$

$$+ \mu_{1j}(x_{ij}) + \gamma_{20}(x_{ij}^2) + \gamma_{21}(z_j)(x_{ij}^2) + \gamma_{22}(m_j)(x_{ij}^2) + \gamma_{23}(z_j)(m_j)(x_{ij}^2) + \mu_{2j}(x_{ij}^2) + r_{ij}$$

Thus, the coefficient for the trial X condition (3 or 4) interaction was γ_{11} and the coefficient for the trial^2 X condition (3 or 4) is γ_{21}. Both interactions were not statistically significant (γ_{11} = -18.06, p = .45; γ_{21} = 1.20, p = .38; see Table 12), indicating that condition did not account for incremental variance beyond that accounted for by the initial change in performance, PPD_2.

These results provide partial support for Hypothesis 4, such that being randomly assigned to the SPD-only condition led to significantly higher performance than being in the GPD-only condition, but the difference became non-significant in the presence of PPD_2. These results indicate that after accounting for the initial change in performance, participants between the two conditions performed no differently.

Table 12

Interactive Effects of Time X Condition (3 v 4) and Time^2 X Condition (3 v 4) Predicting Performance$_{3-8}$

Fixed effects	*b*	*SE*	*df*	*t*	*p*
Time	39.06	16.05	447	2.43	.015
Condition (3 v 4)	26.38	57.46	88	.46	.647
PPD_2	.81	.53	88	1.51	.134
$Time^2$	-2.59	1.52	447	-1.70	.090
Time X Condition (3 v 4)	-18.06	23.95	447	-.75	.451
Time X PPD_2	.01	.22	447	.04	.968
Condition (3 v 4) X PPD_2	.10	.80	88	.12	.905
$Time^2$ X Condition (3 v 4)	2.00	2.28	447	.877	.381
PPD_2 X $Time^2$	.00	.02	447	.04	.967
Time X Condition (3 v 4) X PPD_2	.09	.33	447	.27	.791
$Time^2$ X Condition (3 v 4) X PPD_2	-.01	.03	447	-.30	.767

Discussion

Overview

The purpose of the current study was to examine the effects of over- and underestimations of ability with increasing amounts of task experience. I found consistent results between both studies when predicting subsequent |SPD|s, indicating that people's

perceived ability levels became more accurate over time (i.e., decreasing |SPD|s), which is consistent with prior research finding that past performance is the best predictor of self-efficacy (Atkinson, 1964; Heggestad & Kanfer, 2005). I found that greater initial misjudgments of ability predicted subsequent misjudgments, but contrary to prior research (Baumeister, Smart, & Boden, 1996; Kernis & Johnson, 1990), greater initial misjudgments were unaffected by whether individuals were over- versus underestimating their ability levels. Also, those reporting self-efficacy saw greater improvements in performance than those reporting goal levels. However, the difference between conditions became non-significant when controlling for past-performance (i.e., PPD_2), further suggesting that when predicting future performance, motivational effects are redundant if prior performance data is available. These results contributed to the literature by providing evidence that SPDs have unique motivational effects on skill acquisition and raise issues relating to the usefulness of measuring SPDs, which factors most strongly affect self-efficacy, reporting self-efficacy as opposed to goal levels, and that performance indeed contains a motivational component.

Theoretical and Practical Implications

Usefulness of measuring SPDs. First, this study established a tool for researchers to use when examining the differential effects of self-efficacy on various outcomes. Often, prior research has found inconsistent effects of self-efficacy on performance (Feltz, Chow, & Hepler, 2008; Schmidt & DeShon, 2010; Seo & Ilies, 2009; Vancouver & Kendall, 2006; Vancouver, Thompson, Tischner, & Putka, 2002; Vancouver, Thompson, & Williams, 2001; Yeo & Neal, 2006). Commonly, researchers cite misjudgments of ability level as a reason for conflicting effects of self-efficacy on performance outcomes. The SPD measure constructed for the current study provided specific discrepancy signs and magnitudes which could be used to predict subsequent performance. The measure captures the degree to which individuals misjudge their

own ability levels. Measuring SPDs enables researchers to investigate further the inconsistent effects of self-efficacy on performance and allows trainers to gauge whether learners are over-/underestimating their ability levels. The findings from this study indicate that when considering misjudgments of ability, researchers should use SPDs to consider the magnitude of the initial over-/underestimation because it significantly predicts the development of subsequent over-/underestimations.

Factors affecting self-efficacy. Second, the current study emphasizes the importance of understanding which factors most strongly influence self-efficacy and which factors might cause a non-zero SPD. As shown in the current study, the average |SPD| across participants never became smaller than ~80 points, indicating either (a) that people are poor at predicting their own upcoming performance level or (b) that people might be uninterested in reducing |SPD|s past a certain point. If self-efficacy is primarily a reflection of past performance (Sitzmann & Yeo, 2014), then perhaps self-efficacy is a poorer predictor of subsequent performance than previously thought. Even if participants are using projections in tandem with prior performance when making self-efficacy judgments, the projections would be difficult for participants to gauge with little understanding of learning curves. The decreases in incremental performance in later trials (i.e., performance plateau) would need to be accounted for when making predictive judgements about confidence levels, and the rate would be difficult for a lay-person to predict. Even if participants understood the performance plateau associated with their own learning curves, it would be difficult to gauge where one is on a learning curve. It would require a large amount of attention and critical thinking applied to reporting one's self-efficacy to be able to accurately report how well one expects to perform. Additionally, it is difficult to determine whether participants are poor at predicting their performance later in skill acquisition because of

mental projections or simply because prior performance is progressively getting closer to future performance.

Alternatively, people might be uninterested in reducing |SPD|s past a certain point. According to self-regulatory theories of behavior (Carver & Scheier, 1982; Bandura, 1991), an agent is more likely to exert effort if he perceives a negative discrepancy between current and ideal states, i.e., if actual performance is lower than expected performance. However, if he either (a) doesn't perceive the discrepancy or (b) perceives a *positive* discrepancy, he will not be as motivated to reduce it. Perhaps, like just-noticeable differences in psychophysics, an individual might not perceive a discrepancy unless it is sufficiently large. Individuals might not remember their self-efficacy responses with enough precision to perceive small differences between how they responded and their subsequent performance level. Alternatively, an individual is less likely to reduce a positive- compared to a negative discrepancy because of the distinct emotional reactions. Negative discrepancies (i.e., underperforming relative to one's expectations) leads to unwanted cognitive dissonance. Averting such dissonance would lead one to be more prone to producing positive discrepancies (i.e., overperforming relative to one's expectations). If an individual perceives a positive discrepancy, there is little cognitive dissonance being experienced and consequently little reason for him to reduce his |SPD|, provided he even perceives it in the first place. The current study provided evidence for this possibility such that the mean SPD for each trial after the first was negative (Table 1; note: some might claim the persistent negative SPDs indicates that self-efficacy might be lagging behind performance, such that self-efficacy is simply reflecting prior performance. However, if that was the case, then I would have expected to see |SPD|s approach zero in later trials as changes in performance decreased with the performance plateau). Therefore, |SPD|s might not approach zero because they must be

sufficiently large to be perceived or because individuals underestimating their performance have little reason to reduce the discrepancy.

Reporting self-efficacy versus goal levels. Third, this study provided evidence that reporting self-efficacy before each trial leads to better performance than reporting goal level, which has two possible causes and raises raising distinct issues: effects of rest breaks and effects of goal attainment. One possible cause of differential self-efficacy and goal condition effects is survey length. The self-efficacy measure consisted of 11-items whereas the goal level measure was a single item. Both measures were given before each trial, forcing the participants to stop performing the Class Scheduling Task and reflect on their ability levels. The difference in survey length might have led to those in the self-efficacy condition having a slightly longer rest break between trials compared to those in the goal-level condition (Ariga & Lleras, 2011).

A second possible cause of differential self-efficacy and goal condition effects is goal attainment. That is, those in the goal-level condition reported a single, specific goal level, which might have caused participants to stop trying upon reaching that level. However, those in the self-efficacy condition reported confidence ratings for 11 different performance levels. Reporting confidence for 11 different performance levels might not have provided individuals with a specific target to attain but rather increasingly difficult performance levels to try for as each lower level is attained. This raises the possibility that goal setting might be poorer form of motivational training compared to more ambiguous expectancy scores (i.e., self-efficacy). This is consistent with the weaker effects observed for specific goals when tasks are complex or at low levels of skill acquisition (Locke & Latham, 1990).

Motivational components of performance. Fourth, results from this study emphasize the importance of prior performance when predicting subsequent performance. Condition (SPD-only or GPD-only) significantly predicted differences in performance but became nonsignificant

when PPD_2 was entered in the model, indicating that all the variance in performance accounted for by condition (1 or 2) was subsumed by the variance accounted for by initial differences in performance. This finding illustrates that cognitive and motivational effects of self-efficacy and goal setting might be redundant with or mitigated by information provided by prior performance. The results from the current study are consistent with decades of prior research and theory suggesting that performance is the product of motivation and ability (Anderson & Butzin, 1974; Campbell, 1976).

Although many implications of this study are theoretical, there are several practical implications. Trainers might be interested in whether the trainee is over-/underestimating their ability level throughout the skill acquisition process. The current study provided evidence that the magnitude of the initial discrepancy predicts subsequent discrepancies. Because researchers have cited misjudgments of ability as antecedents to the amounts of resources being allocated to the task (Vancouver & Kendall, 2006; Vancouver, Thompson, Tischner, & Putka, 2002; Vancouver, Thompson, & Williams, 2001), trainers might want to obtain initial discrepancy magnitudes to predict subsequent performance levels. Also, the results of the current study indicate that when training employees/students, if prior performance is not available, having the learner report their self-efficacy before each trial might lead to quicker and greater gains in performance. Reporting self-efficacy might lessen the amount of time required for a learner/trainee to learn new skills in addition to increasing the effectiveness of the training.

Limitations

There were several limitations with the current study that should be addressed. First, the feedback that participants were being given was not as clear as it could have been. The only indication of their performance level was a small section of the screen, and at the end of the trial, the program cut away without explicitly reminding them what their actual point total was.

Perhaps including a screen after the end of each trial stating the cumulative total from that trial would help inform participants' self-efficacy and lead to more accurate assumptions of ability level as we initially predicted. Second, when comparing the effect of condition in Study 2, the current study's design did not include a control condition absent of both goals and self-efficacy. The current study's design used either one or the other, or both. Thus, I do not know what the effects are relative to a no-treatment control condition. Third, when comparing the effects of SPDs and GPDs, the sample size and response formats were not symmetrical between the two conditions. The sample size of Condition 2 was substantially smaller than Condition 1 (Condition 1 $N = 164$, Condition 2 $N = 49$), so the test to detect significant differences between the two conditions might have been underpowered. Also, there were different response formats when collecting SPDs compared to GPDs, such that the self-efficacy measure had 11 items of increasing performance levels whereas the goal level measure was a single self-generated number. Fourth, the current study aimed to capture motivational effects in a sample of undergraduate students who had no perceived consequence for performing poorly. Whereas this creates a more conservative test of motivational effects, it also raises concerns about generalizability. The progression of self-efficacy and SPDs over time might have different effects in high-stakes environments, and the current study failed to capture that potential difference.

Future Research

The results from the current study provide ample directions for future research. First, the current study provided evidence that initial magnitudes of over-/underestimations positively predict subsequent over-/underestimations, indicating a possible stable aspect to misjudgments of ability that is unwavering despite contradictory feedback. People who had large initial misjudgments of ability were more likely to maintain large misjudgments of ability throughout

the subsequent trials. Perhaps some people are simply more likely to be poor predictors of their own ability level, and researchers should investigate which characteristics might explain the distinction. Second, given the limitations of the current sample, future research should investigate whether the findings about SPDs from the current study are consistent in more realistic, high-stakes work environments.

Conclusions

The current study contributes to the existing literature by investigating the progression of misjudgments of ability over time and demonstrating that misjudgments significantly affect performance. Supporting prior research, the current study demonstrated that self-efficacy converges with performance over time (Sitzmann & Yeo, 2013), reporting self-efficacy leads to significantly better performance than reporting goal levels (at least in early skill acquisition; Mitchell et al., 1994; Ordóñez et al., 2009), and inclusion of prior performance renders SPD and GPD effects redundant (Anderson & Butzin, 1974; Campbell, 1976). Future research should investigate the nature of SPDs and their relation to other motivational variables to demonstrate the role of SPDs within the nomological net of motivational variables. In early stages of skill acquisition, trainers should consider asking learners their self-efficacy instead of goal levels before each trial to increase subsequent performance. Overall, these results demonstrate that self-efficacy has a unique role in skill acquisition distinct from goal setting such that self-efficacy – performance discrepancies decrease in magnitude over time and are a valid predictor of subsequent performance.

References

Ariga, A., & Lleras, A. (2011). Brief and rare mental "breaks" keep you focused: Deactivation and reactivation of task goals preempt vigilance decrements. *Cognition, 118*(3), 439-443.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review, 89*(4), 369.

Anderson, N. H., & Butzin, C. A. (1974). Performance= Motivation× Ability: An integration-theoretical analysis. *Journal of Personality and Social Psychology*, *30*(5), 598.

Atkinson, J. W. (1964). An introduction to motivation.

Ashford, S. J., & Tsui, A. S. (1991). Self-regulation for managerial effectiveness: The role of active feedback seeking. *Academy of Management journal*, *34*(2), 251-280.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.

Bandura, A. (1978). The self system in reciprocal determinism. *American psychologist, 33*(4), 344-358.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of social and clinical psychology*, *4*(3), 359-373.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational behavior and human decision processes, 50*(2), 248-287.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan.

Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents, 5*(307-337).

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of personality and social psychology*, *41*(3), 586.

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: the dark side of high self-esteem. *Psychological review*, *103*(1), 5.

Bavelas, J., & Lees, E. S. (1978). Effects of goal level on performance: A trade-off of quantity and quality. *Canadian Journal of Experimental Psychology*, *32*, 219.

Bowers, T. L., & Pantle, M. L. (1998). Shipley institute for living scale and the Kaufman Brief Intelligence Test as screening instruments for intelligence. *Assessment*, *5*(2), 187-195.

Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, *5*(4), 362-387.

Campbell, J. P. (1976). Psychometric theory. *Handbook of industrial and organizational psychology*, 185-222.

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality–social, clinical, and health psychology. *Psychological bulletin*, *92*(1), 111.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational research methods*, *4*(1), 62-83.

Costa, P.T.,Jr. & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual.* Odessa, FL: Psychological Assessment Resources.

Driskell, J. E., Willis, R. P., & Copper, C. (1992). Effect of overlearning on retention.

Emich, K. J. (2014). But consider the alternative: The influence of positive affect on overconfidence. *Cognition and Emotion*, *28*(8), 1382-1397.

Feltz, D. L., Chow, G. M., & Hepler, T. J. (2008). Path analysis of self-efficacy and diving performance revisited. *Journal of Sport and Exercise Psychology*, *30*(3), 401-411.

Fitts, P. M. (1964). Perceptual-motor skill learning. *Categories of human learning*, *47*, 381-391.

Fitts, P. M., & Posner, M. I. (1967). Human performance.

Heggestad, E. D., & Kanfer, R. (2005). The predictive validity of self-efficacy in training performance: little more than past performance. *Journal of Experimental Psychology: Applied*, *11*(2), 84.

Huang, J. L., Bowling, N. A., Liu, M., & Li, Y. (2015). Detecting insufficient effort responding with an infrequency scale: Evaluating validity and participant reactions. *Journal of Business and Psychology*, *30*(2), 299-311.

Ilgen, D., & Davis, C. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology*, *49*(3), 550-565.

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The core self-evaluations scale: Development of a measure. *Personnel psychology*, *56*(2), 303-331.

Kernan, M. C., & Lord, R. G. (1990). Effects of valence, expectancies, and goal-performance discrepancies in single and multiple goal environments. *Journal of applied psychology*, *75*(2), 194.

Kernis, M. H., & Johnson, E. K. (1990). Current and typical self-appraisals: Differential responsiveness to evaluative feedback and implications for emotions. *Journal of Research in Personality*, *24*(2), 241-257.

Krueger, W. F. C. (1930). Further studies in overlearning. *Journal of Experimental Psychology*, *13*(2), 152.

Latham, G. P., & Kinne, S. B. (1974). Improving job performance through training in goal setting. *Journal of Applied Psychology*, *59*(2), 187.

Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational behavior and human performance*, *3*(2), 157-189.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Prentice-Hall, Inc.

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current directions in psychological science*, *15*(5), 265-268.

Marshall, G. N., Wortman, C. B., Kusulas, J. W., Hervig, L. K., & Vickers Jr, R. R. (1992). Distinguishing optimism from pessimism: Relations to fundamental dimensions of mood and personality. *Journal of personality and social psychology*, *62*(6), 1067.

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, *31*(3), 301-306.

Mitchell, T. R., Hopper, H., Daniels, D., George-Falvy, J., & James, L. R. (1994). Predicting self-efficacy and performance during skill acquisition. *Journal of Applied Psychology*, *79*(4), 506.

Moores, T. T., & Chang, J. C. J. (2009). Self-efficacy, overconfidence, and the negative effect on subsequent performance: A field study. *Information & Management*, *46*(2), 69-76.

Niemann, J., Wisse, B., Rus, D., Van Yperen, N. W., & Sassenberg, K. (2014). Anger and attitudinal reactions to negative feedback: The effects of emotional instability and power. *Motivation and Emotion*, *38*(5), 687-699.

Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *Academy of Management Perspectives*, *23*(1), 6-16.

Powers, W. T. (1973). *Behavior: The control of perception* (p. ix). Chicago: Aldine.

Ricciuti, H. N. (1951). A review of procedural variations in level of aspiration studies. *ETS Research Report Series*, *1951*(2).

Richard, E. M., Diefendorff, J. M., & Martin, J. H. (2006). Revisiting the within-person self-efficacy and performance relation. *Human Performance*, *19*(1), 67-87.

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: assessment and implications of generalized outcome expectancies. *Health psychology*, *4*(3), 219.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *Journal of personality and social psychology*, *67*(6), 1063.

Schmidt, A. M., & DeShon, R. P. (2010). The moderating effects of performance ambiguity on the relationship between self-efficacy and performance. *Journal of Applied Psychology*, *95*(3), 572.

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of applied psychology*, *71*(3), 432.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, *84*(1), 1.

Sears, P. S. (1940). Levels of aspiration in academically successful and unsuccessful children. *The Journal of Abnormal and Social Psychology*, *35*(4), 498.

Seijts, G. H., & Latham, G. P. (2001). The effect of distal learning, outcome, and proximal goals on a moderately complex task. *Journal of Organizational Behavior*, 291-307.

Seo, M. G., & Ilies, R. (2009). The role of self-efficacy, goal, and affect in dynamic motivational self-regulation. *Organizational behavior and Human Decision processes*, *109*(2), 120-133.

Shipley, B. (1940). *Shipley Institute of Living scale manual.* Hartford, CT: Neuropsychiatric Press.

Sitzmann, T., & Johnson, S. K. (2012). When is ignorance bliss? The effects of inaccurate self-assessments of knowledge on learning and attrition. *Organizational Behavior and Human Decision Processes*, *117*(1), 192-207.

Sitzmann, T., & Yeo, G. (2013). A Meta-Analytic Investigation of the Within-Person Self-Efficacy Domain: Is Self-Efficacy a Product of Past Performance or a Driver of Future Performance?. *Personnel Psychology*, *66*(3), 531-568.

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

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological bulletin*, *124*(2), 240.

Vance, R. J., & Colella, A. (1990). Effects of two types of feedback on goal acceptance and personal goals. *Journal of Applied Psychology*, *75*(1), 68.

Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology*, *91*(5), 1146.

Vancouver, J. B., Thompson, C. M., & Williams, A. A. (2001). The changing signs in the relationships among self-efficacy, personal goals, and performance. *Journal of Applied Psychology*, *86*(4), 605.

Vancouver, J. B., Thompson, C. M., Tischner, E. C., & Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology*, *87*(3), 506.

Yeo, G. B., & Neal, A. (2006). An examination of the dynamic relationship between self-efficacy and performance across levels of analysis and levels of specificity. *Journal of Applied Psychology*, *91*(5), 1088-1101.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational psychologist*, *25*(1), 3-17.

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American educational research journal*, *29*(3), 663-676.

Appendix A

Demographics

Instructions: The following are a series of demographic questions. Please type the number next to your classification.

1. Sex

   1 = male

   2 = female

   3 = other

2. Age

   1 = 16-18

   2 = 19-21

   3 = 22-24

   4 = 25-27

   5 = 28-30

   6 = 30 and above

3. Year

   1 = Freshman

   2 = Sophomore

   3 = Junior

   4 = Senior

   5 = Post Baccalaureate

   6 = Graduate Student

4. Major

   1 = Business

2 = Communication

3 = Education

4 = Engineering

5 = Mathematics

6 = Psychology

7 = Sociology

8 = Other

5. Overall GPA (Indicate No GPA if you do not have a GPA yet)

1 = 0.0 – 0.5

2 = 0.6 – 1.0

3 = 1.1 – 1.5

4 = 1.6 – 2.0

5 = 2.1 – 2.5

6 = 2.6 – 3.0

7 = 3.1 – 3.5

8 = 3.6 – 4.0

9 = No GPA

6. What is your ethnicity?

1 = White/Caucasian

2 = Black/African American

3 = Asian/Pacific

4 = Hispanic

5 = Native American

6 = Other

7. Are you left handed, right handed, or ambidextrous (can use both left and right hand equally well)?

   1 = Left handed

   2 = Right handed

   3 = Ambidextrous

Appendix B

Task-specific Self-efficacy

Instructions: Please rate your degree of confidence that you can attain the following cumulative point total if you were to complete the Class Scheduling Task again AT THIS MOMENT. (Remember, you earn 20 points for each completed schedule)

Rate your degree of confidence by entering a number from 1 to 9 using the scale given below:

1	2	3	4	5	6	7	8	9
No Confidence				Moderate Confidence				High Confidence

1. How confident are you that you can score at least 0 points (0 Schedules)?
2. How confident are you that you can score at least 20 points (1 Schedule)?
3. How confident are you that you can score at least 40 points (2 Schedules)?
4. How confident are you that you can score at least 80 points (4 Schedules)?
5. How confident are you that you can score at least 120 points (6 Schedules)?
6. How confident are you that you can score at least 160 points (8 Schedules)?
7. How confident are you that you can score at least 200 points (10 Schedules)?
8. How confident are you that you can score at least 240 points (12 Schedules)?
9. How confident are you that you can score at least 280 points (14 Schedules)?
10. How confident are you that you can score at least 320 points (16 Schedules)?
11. How confident are you that you can score at least 360 points (18 Schedules)?

Appendix C

Personality Scale (IPIP NEO-PI-R)

Instructions: Below are phrases describing people's behaviors. Please use the rating scale below to describe how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future.

1	2	3	4	5
Very Inaccurate	Moderately Inaccurate	Neither Inaccurate nor Accurate	Moderately Accurate	Very Accurate

1. Often feel blue. (N)
2. Rarely get irritated. (N)*
3. Dislike myself. (N)
4. Seldom feel blue. (N)*
5. Panic easily. (N)
6. Am often down in the dumps. (N)
7. Feel comfortable with myself. (N)*
8. Am not easily bothered by things. (N)*
9. Have frequent mood swings. (N)
10. Am very pleased with myself. (N)*
11. Feel comfortable around people. (E)
12. Have little to say. (E)*
13. Make friends easily. (E)
14. Keep in the background. (E)*
15. Am skilled in handling social situations. (E)
16. Would describe my experiences as somewhat dull. (E)*
17. Am the life of the party. (E)
18. Don’t like to draw attention to myself. (E)*

19. Know how to captivate people. (E)
20. Don't talk a lot. (E)*
21. Believe in the importance of art. (O)
22. Am not interested in abstract ideas. (O)*
23. Have a vivid imagination (O)
24. Do not like art. (O)*
25. Tend to vote for liberal political candidates. (O)
26. Avoid philosophical discussions. (O)*
27. Carry the conversation to a higher level. (O)
28. Do not enjoy going to art museums. (O)*
29. Enjoy hearing new ideas. (O)
30. Tend to vote for conservative political candidates. (O)*
31. Believe in human goodness. (A-trust)
32. Have a good word for everyone. (A)
33. Make myself the center of attention. (A-modesty)*
34. Have a sharp tongue. (A)*
35. Believe that others have good intentions. (A, A-trust)
36. Cut others to pieces. (A)*
37. Have a high opinion of myself. (A-modesty)*
38. Think that all will be well. (A-trust)
39. Boast about my virtues. (A-modesty)*
40. Respect others. (A)
41. Believe that I am better than others. (A-modesty)*
42. Suspect hidden motives in others. (A, A-trust)*

43. Accept people as they are. (A)
44. Trust what people say. (A-trust)
45. Seldom toot my own horn. (A-modesty)
46. Believe that people are essentially evil. (A-trust)*
47. Get back at others. (A)*
48. Dislike being the center of attention (A-modesty)
49. Believe that people are basically moral. (A-trust)
50. Make people feel at ease. (A)
51. Am wary of others. (A-trust)*
52. Think highly of myself. (A-modesty)*
53. Trust others. (A-trust)
54. Know the answers to many questions. (A-modesty)*
55. Insult people. (A)*
56. Consider myself an average person. (A-modesty)
57. Distrust people. (A-trust)*
58. Dislike talking about myself. (A-modesty)
59. Choose my words with care. (C-cautiousness)
60. Am always prepared. (C)
61. Make rash decisions. (C-cautiousness)*
62. Waste my time. (C)*
63. Pay attention to details. (C)
64. Avoid mistakes. (C-cautiousness)
65. Find it difficult to get down to work. (C)*
66. Jump into things without thinking. (C-cautiousness)*

67. Get chores done right away. (C)

68. Like to act on a whim. (C-cautiousness)*

69. Do just enough work to get by. (C)*

70. Rush into things. (C-cautiousness)*

71. Carry out my plans. (C)

72. Often make last-minute plans. (C-cautiousness)*

73. Don't see things through. (C)*

74. Make plans and stick to them. (C)

75. Do crazy things. (C-cautiousness)*

76. Shirk my duties (C)*

77. Stick to my chosen path. (C-cautiousness)

78. Act without thinking. (C-cautiousness)*

*Reverse scored items. Sub-facet indicated in parentheses.
Scoring: Sum all values of the sub-scale to obtain scores.

Factors:
Neuroticism (N): Alpha = .86
Extraversion (E): Alpha = .86
Openness (O): Alpha = .82
Agreeableness (A): Alpha = .77
Conscientiousness (C): Alpha = .81

Facets:
Cautiousness (C): Alpha = .76
Modesty (A): Alpha = .77
Trust (A): Alpha = .82

Source: International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (http://ipip.ori.org/).

Appendix D

Optimism and Pessimism Scale-Revised (LOT-R)

Instructions: Each of the statements below describes how you might think or feel about challenging life situations. Please answer the questions using the scale below as they relate to your feelings about the challenging situations.

---

1 2 3 4 5

Strongly Disagree Strongly Agree

1. In uncertain times, I usually expect the best.
2. If something can go wrong for me, it will.*
3. I'm always optimistic about my future.
4. I hardly ever expect things to go my way.*
5. I rarely count on good things happening to me.*
6. Overall, I expect more good things to happen to me than bad.

* = *Reverse scored*

Scoring: Average responses; higher scores indicate a higher optimism.

Source: Scheier, Carver, and Bridges (1994)

Appendix E

The Core Self-Evaluations Scale (CSES)

Instructions: Below are several statements about you with which you may agree or disagree. Using the response scale below, indicate your agreement or disagreement with each statement by pressing the appropriate number corresponding to how much you agree or disagree with the statement.

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

1. I am confident I get the success I deserve in life.

2. Sometimes I feel depressed.*

3. When I try, I generally succeed.

4. Sometimes when I fail I feel worthless.*

5. I complete tasks successfully.

6. Sometimes, I do not feel in control of my work.*

7. Overall, I am satisfied with myself.

8. I am filled with doubts about my competence.*

9. I determine what will happen in my life.

10. I do not feel in control of my success in my career.*

11. I am capable of coping with most of my problems.

12. There are times when things look pretty bleak and hopeless to me.*

*Notes: * = reverse-scored.*

Scoring: Average responses; higher scores indicate higher core self-evaluations.

Source: Judge, Erez, Bono, and Thoresen (2003)

Appendix F
Anxiety Scale (STAI – Short Version)

Trait Anxiety

Instructions: A number of statements people use to describe themselves are given below. Read each statement and indicate the extent to which each statement describes *how you generally feel*. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer that best describes your overall feelings best.

1	2	3	4
Not at all			Very Much

1. I feel calm.*
2. I am tense.
3. I feel upset.
4. I am relaxed.*
5. I feel content.*
6. I am worried.

* = *Reverse scored*

Scoring: Average responses; higher scores indicate higher trait anxiety.

Source: Marteau and Bekker (1992)

Appendix G

New General Self-Efficacy Scale

Instructions: Below are statements about people's beliefs that in general they can achieve tasks and goals. Use the following scale to indicate how accurately each item reflects **your own** beliefs about your ability to achieve various tasks and goals.

1	2	3	4	5
Strongly Disagree	Moderately Disagree	Neutral	Moderately Agree	Strongly Agree

1. I will be able to achieve most of the goals that I have set for myself.
2. When facing difficult tasks, I am certain that I will accomplish them.
3. In general, I think that I can obtain outcomes that are important to me.
4. I believe I can succeed at most any endeavor to which I set my mind.
5. I will be able to successfully overcome many challenges.
6. I am confident that I can perform effectively on many different tasks.
7. Compared to other people, I can do most tasks very well.
8. Even when things are tough, I can perform quite well.

Source: Chen, Gully, and Eden (2001)

Scoring: Average responses; higher scores indicate a higher general self-efficacy.

Appendix H

Subjective Effort Scale

Instructions: Each of the statements below describes how you might think or feel about how much effort you exerted on the Class Scheduling Task. Please answer the questions honestly using the given scale for each question.

[1]1. How much effort are you exerting on the current task?

1	2	3	4	5
None		A Moderate Amount		A lot

2. How focused are you on the current task?

1	2	3	4	5
Not at all		A Moderate Amount		A lot

*3. To what extent are you ignoring the instructions for the current task?

1	2	3	4	5
Not at all		A Moderate Amount		A lot

4. How important is it to you to do a good job on the current task?

1	2	3	4	5
Not at all important		Moderately important		Very Important

5. How much attention are you paying to the current task?

1	2	3	4	5
None		A Moderate Amount		A lot

*6. To what extent are you rushing through the current task?

1	2	3	4	5
Not at all		A Moderate Amount		A lot

7. How important is the current task to you?

1	2	3	4	5
Not at all		Moderately		Very

8. How well do you want to perform on the current task?

1	2	3	4	5
Not at all		Moderately		Very

[1]9. How hard are you trying to complete the current task effectively?

1	2	3	4	5
Not at all		Moderately		Very

*10. How difficult is it for you to pay attention to the current task?

1	2	3	4	5
Not at all		Moderately		Very

[1]11. How hard are you trying on the current task?

1	2	3	4	5
Not trying at all		Trying a moderate amount		Trying my best

[1]12. How much effort are you putting into performing the current task?

1	2	3	4	5
None		A Moderate Amount		A lot

13. How motivated are you to complete this task effectively?

1	2	3	4	5
Not at all		Moderately		Very

14. How much energy are you devoting to the current task?

1	2	3	4	5
None		A Moderate Amount		A lot

*15. To what extent are you thinking about things unrelated to the current task?

1	2	3	4	5
Not at all		A Moderate Amount		A lot

[1]16. To what extent are you trying to perform the current task effectively?

1	2	3	4	5
Not at all		Moderately		A lot

17. How much do you care about completing the current task effectively?

1	2	3	4	5
Not at all		Moderately		A lot

18. How carefully are you reading the directions in the current task?

1	2	3	4	5

Not at all Moderately Very

*19. How unmotivated are you to work on the current task?

1	2	3	4	5
Not at all unmotivated		Moderately unmotivated		Very unmotivated

*20. How much attention are you giving to things unrelated to the current task?

1	2	3	4	5
None		A Moderate Amount		A lot

[1] Indicates items retained on the 5-item version.
*Reverse scored items
Scoring: Average scores. Higher scores indicate higher levels of effort exerted on the task.

Revised 5-item version:

Instructions: Each of the statements below describes how you might think or feel about how much effort you exerted on the Class Scheduling Task. Please answer the questions honestly using the given scale for each question.

1. How much effort are you exerting on the current task?

1	2	3	4	5
None		A Moderate Amount		A lot

2. How hard are you trying to complete the current task effectively?

1	2	3	4	5
Not at all		Moderately		Very

3. How hard are you trying on the current task?

1	2	3	4	5
Not trying at all		Trying a moderate amount		Trying my best

4. How much effort are you putting into performing the current task?

1	2	3	4	5
None		A Moderate Amount		A lot

5. To what extent are you trying to perform the current task effectively?

1	2	3	4	5

Not at all Moderately A lot

Scoring: Average scores. Higher scores indicate higher levels of effort exerted on the task.

Appendix I

Cognitive Ability – SILS

Part 1

INSTRUCTIONS: In the test below, the first word in each line is printed in capital letters. Below it are four other words. Shade in the circle next to the word that means the same thing, or most nearly the same thing as the first word. If you do not know, guess. Be sure to shade in the one word in each line that means the same thing as the first word.

1. TALK ○ draw ● speak ○ eat ○ sleep	2. PERMIT ● allow ○ cut ○ sew ○ drive	3. PARDON ● forgive ○ divide ○ pound ○ tell	4. COUCH ○ pin ● sofa ○ eraser ○ glass
5. REMEMBER ○ swim ○ number ● recall ○ defy	6. TUMBLE ○ drink ● fall ○ dress ○ think	7. HIDEOUS ○ silvery ○ young ○ tilted ● dreadful	8. CORDIAL ○ swift ○ leafy ○ muddy ● hearty
9. EVIDENT ○ green ○ skeptical ● obvious ○ afraid	10. IMPOSTOR ○ conductor ○ book ○ officer ● pretender	11. MERIT ● deserve ○ fight ○ distrust ○ separate	12. FASCINATE ○ welcome ○ stir ○ fix ● enchant
13. INDICATE ○ defy ● signify ○ excite ○ bicker	14. IGNORANT ○ red ● uninformed ○ sharp ○ precise	15. FORTIFY ○ submerge ○ vent ● strengthen ○ deaden	16. RENOWN ○ length ● fame ○ head ○ loyalty
17. NARRATE ○ yield ○ associate ○ buy ● tell	18. MASSIVE ○ bright ○ speedy ● large ○ low	19. HILARITY ● laughter ○ grace ○ speed ○ malice	20. SMIRCHED ○ stolen ○ remade ○ pointed ● soiled
21. SQUANDER ○ tease ○ cut ○ belittle ● waster	22. CAPTION ○ drum ● heading ○ ballast ○ ape	23. FACILITATE ● help ○ strip ○ turn ○ bewilder	24. JOCOSE ● humorous ○ fervid ○ paltry ○ plain

25. APPRISE ○ reduce ● inform ○ strew ○ delight	**26. RUE** ○ eat ○ dominate ● lament ○ cure	**27. DENIZEN** ○ senator ○ fish ● inhabitant ○ atom	**28. DIVEST** ● dispossess ○ rally ○ intrude ○ pledge
29. AMULET ● charm ○ dingo ○ orphan ○ pond	**30. INEXORABLE** ○ untidy ● rigid ○ involatile ○ sparse	**31. SERRATED** ○ dried ○ armed ● notched ○ blunt	**32. LISSOM** ○ moldy ● supple ○ loose ○ convex
33. MOLLIFY ● mitigate ○ pertain ○ direct ○ abuse	**34. PLAGIARIZE** ● appropriate ○ revoke ○ intend ○ maintain	**35. ORIFICE** ○ brush ○ building ● hole ○ lute	**36. QUERULOUS** ○ maniacal ○ devout ○ curious ● complaining
37. PARIAH ● outcast ○ lentil ○ priest ○ locker	**38. ABET** ○ waken ● incite ○ ensue ○ placate	**39. TEMERITY** ● rashness ○ desire ○ timidity ○ kindness	**40. PRISTINE** ○ vain ● first ○ sound ○ loyalty

Part 2

INSTRUCTIONS: Complete the following by filling in either a number or a letter for each dash (___). Do the items in order, but do not spend too much time on any one item.

1. 1 2 3 4 5 _6_
2. white black short long down _U P_
3. AB BC CD D _E_
4. Z Y X W V U _T_
5. 1 2 3 2 1 2 3 4 3 2 3 4 5 4 3 4 5 6 _5 4_
6. NE/SW SE/NW E/W N/ _S_
7. escape scape cape _A P E_
8. oh ho rat tar mood _D O O M_
9. A Z B Y C X D _W_
10. tot tot bard drab 537 _7 3 5_
11. mist is wasp as pint in tone _O N_
12. 57326 73265 32657 26573 _6 5 7 3 2_
13. knit in spud up both to stay _A T_
14. Scotland landscape scapegoat _G O A T_ ee
15. surgeon 1234567 snore 17635 rogue _3 6 4 2 5_
16. tam tan rib rid rat raw hip _H I T_
17. tar pitch throw saloon bar rod fee tip end plank _B O A R D_ meals
18. 3124 82 73 154 46 13 _6_
19. lag leg pen pin big bog rob _R U B_
20. two w four r one o three _R_

Shipley, B. (1940). *Shipley Institute of Living scale manual.* Hartford, CT: Neuropsychiatric Press.

Appendix J

Consent form

Subject No._____

Subject Informed Consent Document

LEARNING EFFECTS OF A CLASS SCHEDULING TASK

Investigator(s) name: Kent Etherton (Psychology Dept.)
Site(s) where study is to be conducted: Wright State University
Email for subjects to contact for questions: Etherton.4@wright.edu

Introduction and Background Information

You are invited to participate in a research study. The study is being conducted by Kent Etherton under the supervision of Dr. Debra Steele-Johnson. Approximately 300 subjects will be invited to participate.

Purpose

The purpose of this research is to (a) examine how people manage perceived self-efficacy – performance discrepancies during a series of trials and (b) what effect the discrepancies have on subsequent effort and performance. Examining the role of self-efficacy – performance discrepancies will aid the understanding of how over- and underconfident self-efficacy judgements progress with experience.

Procedure

In this study, you will be asked to complete eight ten-minute trials of a computerized task. The task will involve you creating as many academic schedules as possible without violating a set of rules which will be described to you. Also, you will be asked to complete questionnaires assessing demographic information, effort, personality, and self-efficacy. Demographic information will be collected to examine possible differences in performance between genders, ethnicities, ages, etc. The expected time to complete this study is 3 hours. If you do not complete the study in 3 hours, you will be asked to stay and complete the study, but you will not be penalized in any way for leaving at that time. If you leave at the end of the 3 hours, you will still receive full SONA credits.

Potential Risks

There are no foreseeable risks other than fatigue and possible discomfort in answering personal questions.

Benefits

There are no anticipated benefits to you for participating in this study. The information learned in this study may be helpful to others.

Compensation

You will receive 6 research credits through SONA for your participation in this study. If you leave early, you will receive 1 research credit for every 30 minutes you participated in the study.

Confidentiality

Total privacy cannot be guaranteed. We will protect your privacy to the extent permitted by law. If the results from this study are published, your name will not be made public. Once your information leaves our institution, we cannot promise that others will keep it private.

Your information may be shared with the following:

- The Wright State IRB and Office of Research and Sponsored Programs
- Office for Human Research Protections (OHRP),

Security

The consent form, the only source of your name, will be kept securely in the principal investigator's locked office. The data collected from the study will be secured on a password protected computer.

Voluntary Participation

Taking part in this study is voluntary. You may choose not to take part at all. If you decide to be in this study you may stop taking part at any time. However, if you decide not to be in this study or if you stop taking part at any time, you will no longer qualify to receive research credits through SONA.

Research Subject's Rights, Questions, Concerns, and Complaints

You may contact the principal investigator, Kent Etherton, at Etherton.4@wright.edu or his faculty advisor, Dr. Debra Steele-Johnson, at debra.steele-johnson@wright.edu.

If you have any questions about your rights as a study subject, questions, concerns, or complaints, you may call the Wright State IRB Office (937) 775-4462. You may discuss any questions about your rights as a subject with a member of the IRB or staff. The IRB is an independent committee composed of members of the University community, staff of the institution, as well as lay members of the community not connected with the institution. The IRB has reviewed this study.

This paper tells you what will happen during the study if you choose to take part. Your signature means that this study has been discussed with you, that your questions have been answered, and that you will take part in the study. This informed consent document is not a contract. You are

not giving up any legal rights by signing this informed consent document. You will be given a signed copy of this consent to keep for your records.

Signature of Subject	**Date Signed**
Signature of Person Obtaining Consent (if other than the Investigator)	**Date Signed**
Signature of Investigator	**Date Signed**

Appendix K

Debriefing form

The study you just completed is concerned with examining your performance on and perceptions of a computerized class scheduling task. We have asked you to create class schedules on several trials of the computerized task and complete multiple surveys that will give us information about your levels of confidence regarding the task.

<Press the SPACE BAR to move to the next section.>

Please do not discuss these questionnaires or the computer task with anyone else, for it is important that future subjects know nothing about the study before they begin it.

The data you provided today is important to us, and we appreciate your help. If you have any questions or comments about today's study, please talk to the researcher now or contact Kent Etherton (317) 450-7620 or Debra Steele-Johnson, Ph.D., Department of Psychology, 325B Fawcett, (937) 775-3527.

Thank you for your assistance. The session will end when you press the space bar.

<Press the SPACE BAR to move to the next section.>

The end