2009

Error Management Training: Further Tests of Mediation and Moderation

Zachary T. Kalinoski
Wright State University

Follow this and additional works at: https://corescholar.libraries.wright.edu/etd_all

Part of the Industrial and Organizational Psychology Commons

Repository Citation
https://corescholar.libraries.wright.edu/etd_all/302

This Thesis is brought to you for free and open access by the Theses and Dissertations at CORE Scholar. It has been accepted for inclusion in Browse all Theses and Dissertations by an authorized administrator of CORE Scholar. For more information, please contact corescholar@www.libraries.wright.edu, library-corescholar@wright.edu.
ERROR MANAGEMENT TRAINING: FURTHER TESTS OF MEDIATION AND MODERATION

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science

By

ZACH KALINOSKI
B. A., Ohio University, 2007

2009
Wright State University

Debra Steele-Johnson, Ph.D.
Thesis Director

Dan Weber, Ph.D.
Graduate Program Director

John Flach, Ph.D.
Chair, Department of Psychology

Committee on Final Examination

Debra Steele-Johnson, Ph.D.

Nathan Bowling, Ph.D.

Dragana Claflin, Ph.D

Joseph F. Thomas, Jr., Ph.D.
Dean, School of Graduate Studies
ABSTRACT


This study investigated an alternative training approach that would improve transfer performance scores above traditional training approaches. Specifically, error-management training was proposed to help trainees learn complex tasks, as opposed to error-avoidant training approaches, which sought to give trainees step-by-step protocols for learning that would minimize the occurrence of errors during training. This study was designed to examine the effects of training type on transfer performance and transfer errors, as well as the effects of meta-cognition, emotion control and cognitive appraisals as mediators of the training type-performance relationship. A third issue of this study investigated the personality-training type interactions from a situation strength perspective. Participants (N = 181) from a Midwestern university completed four training trials and two transfer trials of a computerized version of a class scheduling task and completed surveys of relevant constructs. Results revealed that training type did not have an effect on transfer performance or errors, training type did not predict meta-cognition, emotion control and challenge appraisals, but did predict threat appraisals. Finally, personality did not have a main effect on performance, nor did it interact with training type. The relative contributions of this study was the effects of training type on cognitive appraisals (threat in particular) and its relevance for future theoretical frameworks of error management training research, the effects of training type on error attitudes and error attitude effects on performance. Previous operationalizations of error management training also may not be as clear-cut as once thought.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.  INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Encouraging Errors in Training</td>
<td>3</td>
</tr>
<tr>
<td>Mechanisms of Error Management Training</td>
<td>6</td>
</tr>
<tr>
<td>Meta-Cognition</td>
<td>7</td>
</tr>
<tr>
<td>Emotion Control</td>
<td>7</td>
</tr>
<tr>
<td>Cognitive Appraisals</td>
<td>8</td>
</tr>
<tr>
<td>Moderation of Error Management Training</td>
<td>11</td>
</tr>
<tr>
<td>Past Research on Moderators</td>
<td>11</td>
</tr>
<tr>
<td>Role of Situation Strength in Training Type Effects</td>
<td>13</td>
</tr>
<tr>
<td>Perfectionism</td>
<td>17</td>
</tr>
<tr>
<td>Optimism</td>
<td>19</td>
</tr>
<tr>
<td>II. METHOD</td>
<td>22</td>
</tr>
<tr>
<td>Participants</td>
<td>22</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>22</td>
</tr>
<tr>
<td>Task Description</td>
<td>22</td>
</tr>
<tr>
<td>Manipulation Description</td>
<td>24</td>
</tr>
<tr>
<td>Measures</td>
<td>26</td>
</tr>
<tr>
<td>Procedure</td>
<td>31</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>34</td>
</tr>
<tr>
<td>Sample Characteristics</td>
<td>34</td>
</tr>
<tr>
<td>Sample Size</td>
<td>34</td>
</tr>
<tr>
<td>Initial Group Differences</td>
<td>34</td>
</tr>
<tr>
<td>Manipulation Check</td>
<td>37</td>
</tr>
<tr>
<td>Psychometrics</td>
<td>37</td>
</tr>
<tr>
<td>Transfer Task Operationalization</td>
<td>39</td>
</tr>
<tr>
<td>Manipulation Check by Condition</td>
<td>39</td>
</tr>
<tr>
<td>Tests of Hypotheses</td>
<td>42</td>
</tr>
<tr>
<td>IV. DISCUSSION</td>
<td>53</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Framing Statement Heuristics</td>
<td>25</td>
</tr>
<tr>
<td>2. Manipulation Components by Condition</td>
<td>26</td>
</tr>
<tr>
<td>3. Experimental Procedure Outline</td>
<td>33</td>
</tr>
<tr>
<td>4. Correlation Matrix</td>
<td>35</td>
</tr>
<tr>
<td>5. Factor Analysis of Manipulation Check Measure</td>
<td>41</td>
</tr>
<tr>
<td>6. Manipulation Check Means and Standard Deviations by Condition</td>
<td>42</td>
</tr>
<tr>
<td>7. Means and Standard Deviations by Condition on Various Outcomes</td>
<td>45</td>
</tr>
<tr>
<td>8. Summary of Main Effects and Mediation Analyses on Transfer Schedules</td>
<td>46</td>
</tr>
<tr>
<td>9. Summary of Main Effects and Mediation Analyses on Transfer Errors</td>
<td>47</td>
</tr>
<tr>
<td>10. Summary of Moderation Analyses for Adaptive Perfectionism</td>
<td>49</td>
</tr>
<tr>
<td>11. Summary of Moderation Analyses for Maladaptive Perfectionism</td>
<td>50</td>
</tr>
<tr>
<td>12. Summary of Moderation Analyses for Optimism</td>
<td>52</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Conceptual Model</td>
<td>21</td>
</tr>
</tbody>
</table>
Acknowledgement

I would like to thank Dr. Debra Steele-Johnson, my advisor, for her constant support and oversight of this project. Her continual attempts to revise and improve this thesis were commendable and much appreciated. She always encourages her students to think of new topic areas that were of particular interest to the students, which is refreshing. Her commitment to these new ideas during the process of the thesis was second-to-none.

I also would like to thank my other two committee members, Dr. Nathan Bowling and Dr. Dragana Claflin. These two always kept Deb and I on our toes and showed us sides of the project that we could not see ourselves. Their support was much appreciated.

I would like to commend the members of my lab for their insight the past two years. Keith Leas was of tremendous help, as were the two other senior students, Anupama Narayan and Kristin Delgado. I would also like to extend my gratitude to Julie Steinke and Dorothy Carter for their assistance as well.

Not to mention, I want to give special mention to my I/O cohort, Erin O’Brien, Cristina Kirkendall, and Jenna Filipkowski. All of those times discussing this project in Research Methods helped me to define this project. I thank you all once again for your insight and comments.

Thank you to Dr. Michael Frese, for your time in Colorado for discussing error management training and control theory with me. Our hour-long conversation was very useful in defining a portion of this project.

Finally, I wish to express my sincerest of gratitude to the loves of my life, my parents and brother, for all of the familial love over the years. Their commitment to me is unconditional and I am forever in their debt. My wonderful girlfriend, Alison, who has been there for me since day one, has shown undying commitment to me, especially during grad school. Thank you for tolerating my late-night philosophical rants over the phone! To all my friends and extended relatives, thanks again! You know you all mean the world to me. I appreciate your support.
Introduction

Organizations are changing the ways in which they operate (Cascio, 1995). As organizations become more horizontal in design and face rapid technological changes in the global market, their workforces must be equipped with the proper knowledge and skills to enable change at a quick pace. Weick and Quinn (1999) proposed a continuous change model for designing organizations that create horizontal structuring and loosening departmental interdependencies while empowering smaller units to fix problems that may arise. An organization adopting the continuous change philosophy places great value on learning, improvisation, and development. Affording employees the opportunity to learn and progress within their jobs has a twofold benefit: furthering organizational and individual goals. Training plays a large part in satisfying both goals. Individually, training can equip employees with knowledge and skills that can enhance career success (Kim, 2005). Further, training has been shown to increase organizational effectiveness and gain market share by increasing a firm’s intellectual capital (Noe, 1999; Salas & Cannon-Bowers, 2001). Unfortunately, the traditional types of training that are so common in organizations are out-dated and need revision to fit in with the new organizational model.

Traditionally, most training interventions have been formal, proceduralized means of providing information to employees. For instance, the most common training method used by organizations is classroom instruction (e.g., lecture; Noe, 1999). Although there are advantages to this method (e.g., mass dissemination of information), the primary weakness of this approach is that it encourages passive participation, rote memorization, and an error-avoidant attitude on behalf of participants. This approach has been called error-avoidant training, and it discourages
trainees from making mistakes (Frese, Broadbeck, Heinbokel, Mooser, Schleiffenbaum, & Thiemann, 1991).

Keith and Frese (2008) demonstrated that, across trials, the different training systems have differential effects on performance. Specifically, error avoidant training would be more beneficial early on in skill acquisition whereas the error management training condition would have a detrimental effect on performance. Error avoidant training provides early guided mastery practice when the task is novel to the participant whereas error management training would encourage exploratory behavior, leading to a greater amount of errors early in the skill acquisition process. These errors may, in turn, cause individuals to further explore the task and test hypotheses more than if mistakes were not committed, thus leading to a refinement of previous task schematization (Dormann & Frese, 1994). This exploration would increase the amount of declarative knowledge (Bell & Kozlowski, 2008) that would not have existed before the commitment of errors occurred.

Keith and Frese (2008) conceded that in certain situations, particularly when the training task is simple and analogical to the real job setting, error avoidant training is just as effective as any alternative. However, when the task is novel or complex, a different kind of training can have greater effects on transfer performance and learning, such as error-management training. Training sessions are limited in how much information can be taught. Therefore, errors are bound to occur, both in training and on the job. Teaching trainees how to handle task ambiguities and errors in a productive manner can aid organizations modeled to adapt to continuous change. These organizations require competent, self-reliant workers who can address problems as they are encountered on the job. Often, these problems are complex and
unpredictable and require novel thinking and problem-solving strategies. Error management training helps to overcome some of the deficiencies of traditional training.

*Encouraging Errors in Training*

Error training represents the broad philosophy that errors are beneficial when learning and has taken on many forms in previous research (error management training being one example). We will discuss three common types of error training. First, Lazar and Norcio (2003) used an *exploratory learning* orientation and defined it as “encountering objects and situations with a degree of uncertainty.” Exploratory learning encourages active learning on behalf of the participant and implies very little structure in which to perform the task. Exploration is important for task performance, but depending on the particular strategy used by participants (i.e., systematic or unsystematic strategies), exploration can have differential effects (van der Linden, Sonnentag, Frese, & van Dyck, 2001). With systematic exploration strategies, individuals engage in hypothesis testing and refine their strategies based on feedback that points to the success or failure of previous task attempts. With unsystematic exploratory strategies, one can engage in trial-and-error, rigid exploration, or encapsulation in information-seeking. On the whole, van der Linden, et al. (2001) discovered that systematic exploration led to more positive consequences (e.g., task performance, positive consequences from errors, etc.) for individuals than the unsystematic strategies.

Another form of error training is called *enactive exploration*. The intention of this type of training is to give trainees the opportunity to familiarize themselves with the task initially and gain a basic level of competence before exploring the task. Wood, Kakebeeke, Debowski and Frese (2000) showed that this approach increased performance, satisfaction, and self-efficacy with the task relative to guided mastery training alone.
Error-management training involves two components: (1.) exploration of the task and (2.) positive framing statements (Keith & Frese, 2008). Framing statements (e.g., “Mistakes are a natural part of learning,” “Do not be afraid of making errors”) are intended to mitigate the negative reactions associated with personal errors or mistakes. Negative arousal and mistakes are not prevented from occurring by framing statements but rather are re-framed cognitively so as to maximize their benefits for learning (Nordstrom et al., 1998).

Keith and Frese (2008) conducted a meta-analysis on the available literature and found that, overall, error management training resulted in higher transfer performance scores than error avoidant training ($d = .44$). Their research also pointed towards the acceptance of error management training over mere exploratory learning because the moderation analysis demonstrated that exploration plus framing statements resulted in greater transfer performance scores than exploratory learning alone ($d = .19$). Given this knowledge, it is reasonable to look at some possible reasons for the negative results observed by some researchers (Gully et al., 2002; Lazar & Norcio, 2003).

For instance, Gully et al. (2002), when assessing performance, actually measured training performance and not transfer performance. When assessing training performance, researchers should expect inferior performance scores because participants are still learning the task and make more errors in training performance. On a structurally distinct task (i.e., transfer performance task), error management training participants are expected to perform better (Keith & Frese, 2005; 2008). These tasks generally take on the form of increased difficulty and are recognizably different than any task experienced in training. Thus, participants trained in an error avoidant training condition no longer can rely on the crutch of step-by-step instructions and must adapt to the new demands of the task. However, error management training participants are
better equipped to handle such situations because in a benign environment errors and exploration forced participants to refine their strategy and problem-solving development, and this resulted in a better grasp of the deep structure schema that underlies the task. The negative relationship between error management training and performance in Gully et al. (2002) can be attributed to improper operationalization of performance.

Lazar and Norcio (2003), on the other hand, conceptualized error management training differently than the definition proffered by Keith and Frese (2008). These researchers operationalized error management training as error avoidant training plus framing statements, disregarding the exploratory component. Lazar and Norcio’s (2003) error management training concept was the least effective training type in comparison to the other training types, with exploratory training resulting in the highest performance score. This pattern might be an artifact of the task (i.e., Internet search task). The researchers also believed that counting errors was irrelevant because the error training methodology viewed errors as a good thing. Taking notice of errors would have proved useful considering error management training is supposed to induce more errors in a training session than error avoidant training. If participants do not commit more errors than error avoidant training participants, then the manipulation might not have been strong enough to enact the desired psychological effect.

Finally, Debowski, Wood, & Bandura (2001) found that guided exploration participants formed better strategies, had higher performance, self-efficacy, and satisfaction and wasted less effort than enactive exploration participants. Despite these findings, Debowski et al. (2001) said that guided exploration may have been more congruent with the nature of the electronic search task because task feedback did not provide useful information as to whether a particular strategy was successful or not. Enactive exploration requires useful task feedback in order to learn and
refine strategies. Guided exploration is predetermined and systematic in the sense that modes of exploration are bestowed upon trainees as opposed to being self-regulated or self-produced. As Keith and Frese (2008) stated, error management training (more akin to enactive exploration than guided exploration) is more effective when the task provides clear task feedback than when it provides ambiguous or no feedback. Thus, the Debowski et al. (2001) finding is not surprising given our current knowledge of error-management training practices.

Despite the fact that the above studies that have shown negative results for error management training on performance (Debowski, Wood, & Bandura, 2001; Gully, Payne, Koles, & Whiteman, 2002; Lazar & Norcio, 2003), the majority have revealed positive effects. For instance, error management training has increased post-training performance scores (Chillarege, Nordstrom, & Williams, 2003; Frese et al., 1991; Keith & Frese, 2005; Nordstrom, Wenland, & Williams, 1998; Wood, Kakebeeke, Debowski, & Frese, 2000), requests for assistance (Chillarege et al., 2003), and intrinsic motivation levels (Nordstrom et al., 1998; Wood et al., 2000) and decreased frustration levels (Nordstrom et al., 1998). Error management training coupled with state learning goals also has shown to improve performance as well (Nordstrom et al., 1998). Thus, I expect to replicate previous results.

H1: Training type will affect transfer performance. Error management training participants will demonstrate higher performance than error avoidant training participants.

Mechanisms of Error Management Training

Keith and Frese (2005) empirically tested possible mediating mechanisms that underlie error management training effectiveness and found that trainees undergoing error management training engaged in more metacognitive and emotional control activities than those undergoing error avoidant training. In essence, error management trainees put more thought into the task
and were better able to control their negative emotional arousal resulting from the errors made during the training session.

Metacognition. Making mistakes allows individuals to engage in deeper level processes to better grasp the nature of the task (Fisher & Lipson, 1986). Errors as witnessed in a training session give trainees the opportunity to explore the task in a way not available if mistakes were minimal or discouraged from occurring. In error management training, trainees will develop better mental models and learning strategies than individuals participating in an error avoidant training session. By thinking over mistakes under the assumption that errors form learning opportunities, one is taught to refine previous strategies that may or may have not worked. Deeper levels of processing refer to a “greater degree of semantic or cognitive analysis” (Craik & Lockhart, 1972). Craik and Lockhart (1972) also suggested that retaining information is a function of depth of processing, which depends on the amount of attention given to a stimulus and the processing time available. Deeper levels of processing also imply more use of past knowledge and rules, thus aiding the transfer of learning to the task setting.

H2: Meta-cognition will partially mediate the effect of training type on transfer performance such that error management training participants will employ more meta-cognitive techniques which will lead to higher performance, relative to error avoidant training participants.

Emotion control. Learning may be inhibited by either deterring the occurrence of or the individual processing of mistakes. Ohlsson (1996) proposed that in order for individuals to learn from performance errors, one must first detect that the error occurred, and second, correct the error via identified knowledge reconstruction. However, individuals would prefer to gravitate towards error-avoidant behaviors because individuals perceive that mistakes reflect badly on the
individual and these perceptions result in negative emotional arousal (Hajcak & Foti, 2008). The idea behind error management training is to allow individuals the opportunity to make mistakes in a training session and keep their focus on the task and away from the self, considering individuals perceive more feedback from the self than the task or external source (Ilgen, Fisher, & Taylor, 1979). In error management training, the training session is structured so that the trainee receives encouragement to explore the task and learn from errors. Positive heuristics, or framing statements, are used to minimize the negative arousal associated with the mistakes made in a training session. Error avoidant training, on the other hand, fosters step-by-step procedures for individuals to abide by, in addition to the aid of the trainer when mistakes are made. The latter reduces the likelihood that trainees will experience the error and instead are given the solution directly. This may benefit the trainee within the training session, but have a negative impact on transfer performance because they were unable to refine their strategies within training, thus reducing the potential to formulate a richer mental model of the task.

H3: Emotion control will partially mediate the effect of training type on transfer performance such that error management training participants will employ more emotional control techniques which will lead to higher performance, relative to error avoidant training participants.

Cognitive appraisals. As the above discussion suggested, the positive effects of error management training over error avoidant training on performance can be accounted for by metacognitive and emotion control activities on behalf of the participant. Perhaps an additional explanatory mechanism at play in this process is the role of cognitive appraisals. Cognitive appraisals originated in the stress and coping literature and has been shown to relate to variables relevant in the work domain, such as task performance (Skinner & Brewer, 2002), and self-
efficacy (Boggiano, Main & Katz, 1988; Karademas & Kalantzi-Azizi, 2004; Major, Richards, Cooper, Cozzarelli, & Zubek, 1998). Considering that errors invoke stress-inducing cognitions and affect, it is appropriate for error management training researchers to adopt a theoretical perspective that can account for the thought processes that precede stressful events in a training context.

Cognitive appraisals, based on Lazarus and Folkman’s (1984) original model, denote “person/situation-evoked motivational states that include affective (or emotional), cognitive, and physiological components” (Blascovich & Mendes, 2000). More importantly, primary (i.e., task demands) and secondary (i.e., personal resources) appraisals are made by individuals when faced with a task or situation that needs rectified or performed. These appraisals can induce three separate stress responses: harm/loss (which has already been accrued), threat, and challenge (which are both anticipatory). My focus is on the anticipatory appraisals. If primary demands exceed secondary appraisal resources, there is deemed a potential for harm or loss. Thus, the person would tend to view the task as threatening, leading to fatalistic thinking (McCrae, 1984), anxiety and decreased performance (Skinner & Brewer, 2002). If threat appraisals pre-dominate one’s functioning, these particular individuals will tend to avoid stress-induced situations and turn in-ward (Lazarus & Folkman, 1984). Ilgen et al. (1979) and Hajcak and Foti (2008) would lend support to the idea of inwardness as it relates to feedback perception and negative arousal, respectively.

If, on the other hand, resources are sufficient to handle the demands of the task and there is deemed a potential for mastery or gain, a cognitive appraisal of challenge is thought to arise, evoking emotions and thoughts such as positive, rational thinking, and self-adaptation (McCrae,
Physiological stress responses also differ depending on if one feels threatened or challenged by activities (Blascovich & Mendes, 2000; Lazarus & Folkman, 1984).

Another intriguing aspect of cognitive appraisals is the reappraisal process. Lazarus and Folkman (1984) suggested that an individual can experience ever-changing appraisals throughout the process of task engagement with the time elapse varying between a few minutes to a few days. Depending on the changing nature of the task, a task can begin as challenging but be redefined as threatening as individuals assess their performance. This reappraisal process bears relevance within the training literature, specifically in regards to error management training and error avoidant training interventions, because the training environment is a dynamic, ever-changing learning process and would foster reappraisals.

Error management training participants tend to exhibit higher levels of intrinsic motivation than error avoidant training participants on the adaptive transfer trials (Nordstrom et al., 1998; Wood, et al., 2000). As a reminder, a transfer trial will assess trainees’ ability to adapt and use skills learned in training to either similar tasks or structurally distinct activities (sometimes referred to as near and far transfer). Higher levels of intrinsic interest in a task may lead to preference for challenging activities (Boggiano, Main, & Katz, 1988; Inoue, 2007). As competence and skill is attained, levels of interest and task enjoyment may rise, thus influencing appraisals and perceptions of challenge. Participants also tend to experience decreased frustration levels, in comparison to error avoidant training participants (Nordstrom et al., 1998). I expect that framing errors as positive and beneficial to participants may alter the manner in which they perceive the task as threatening or challenging. Because the framing statements and all-around induction are intended to mitigate negative thoughts and emotions, I would expect error management training participants to view the task as more challenging. Error avoidant
training participants, on the other hand, tend to perform worse on transfer trials (Keith & Frese, 2008) and experience higher frustration levels (Nordstrom et al., 1998).

H₄a: Cognitive appraisals will partially mediate training type effects on transfer performance such that error management training participants will appraise the task as more challenging which will lead to higher performance, relative to error avoidant training participants.

H₄b: Cognitive appraisals will partially mediate training type effects on transfer performance such that error avoidant training participants will appraise the task as more threatening which will lead to lower performance, relative to error management training participants.

Moderation of Error Management Training

Error management training researchers have noted several moderators of the relationship between training type and performance. For instance, Gully, Payne, Koles, and Whiteman (2002) reported that individuals high in cognitive ability benefit from an error management training intervention more so than low cognitive ability individuals. Keith and Frese (2008) suggested that the effects of training type on transfer performance are determined by the transfer task type. In other words, analogical transfer tasks within the job setting would resemble the task as experienced within a training context. Given this synchronicity between settings, the authors stated that either error management training or error avoidant training would predict performance equally. Adaptive transfer tasks, on the other hand, involve structurally distinct tasks from those learned within the training setting. These types of tasks prove more challenging for the individual and provide a more stringent test of the strategies learned within training. Here, error management training resulted in higher transfer performance scores than error avoidant training (d = .80). Finally, Keith and Frese (2008) found that task feedback moderates the effects of error
management training on performance. If the task does not provide clear feedback, exploration would prove futile because feedback helps reduce goal discrepancies (Frese & Zapf, 1994). In the case of no task feedback, error avoidant training would be more beneficial because of the intervention’s reliance on external feedback (Debowski et al., 2001). Much of the early feedback research dealt with task feedback (or “knowledge of results”) and concluded that feedback does improve performance (both speed and level of learning; Ammons, 1956; Arps, 1920; Judd, 1905). Specific and immediate feedback also tends to improve performance (Ammons, 1956).

One question that arises concerns the generalizability of error management training research to different types of jobs. Most research has focused upon various computer tasks (e.g., Word, Excel, PowerPoint, SPSS, etc.) or decision-making tasks (i.e., TANDEMS, radar-tracking). Jobs high in computer proficiency as well as naval radar-trackers would benefit from error management training. Some jobs have grave consequences if errors occur. This would indicate that error management training may not be applicable to all jobs. However, some researchers have looked at high-risk jobs using other tactics such as simulations and vicarious learning.

Ivancic & Hesketh (2000) investigated the use of error exposure in driver training simulation, compared to error-less simulation training. Ivancic & Hesketh (2000) operationalized errors as speeding tickets, crashing into concrete barriers, etc. Thus, trainees experienced the errors but did not suffer the negative consequences that would come about from experiencing these errors in a real-world scenario. Exploration of errors was not possible in this study. The authors also assessed guided versions of the training, which had less of an effect than the driving simulation.
Joung, Hesketh, & Neal (2006) used vicarious learning to teach firefighters crisis management. To do so, they compared two groups, one receiving success stories and the other receiving failure stories about handling past crisis events. Trainees learned better from the mistakes of others than from the successes of others.

Finally, Rogers, Regher, & MacDonald (1972) investigated the vicarious effects of learning surgical skills (i.e., tying a knot). Specifically, they compared a control group with a group receiving correct video instructions, a group receiving incorrect video instructions and finally a group receiving both sets of instructions. Trainees shown both the correct and the incorrect way to tie a knot outperformed the other groups.

Thus, it remains to be seen the extent of jobs that may benefit from error management training. Even if the consequences of making errors in training are grave (as would be the case in surgical training or fire-fighting), previous research have found ways, via simulations and vicarious learning, to circumvent the negative ramifications that may arise from errors. This increases the extent to which error management training can be applied. Also, we may not know the extent to which particular jobs, per sé, are most appropriate for error management training, but rather what types of skills that need to be trained. For instance, error avoidant training would be appropriate for simple skills with little variation whereas complex skills may benefit from error management training more so than error avoidant training.

What remains unclear in past research, which is a major focus of this study, is the effect of personality on the relationships between training type and performance. Inconsistent results have prevented researchers from reaching conclusions about personality as a potential moderator (Gully et al., 2002; Heimbeck et al., 2003). We offer an alternative perspective to guide future investigations of the training type-personality interaction effects on performance.
The role of situation strength in training type effects. The strong/weak situation distinction was first introduced by Mischel (1968), who called for a more dynamic approach to studying human personality. Mischel (1968) found the personality conceptualizations of the time to be unsatisfactory because he deemed the personality theories as too rigid and believed individuals were “victimized by [their] infantile history.” In other words, individuals were construed and understood as genetically predisposed to behave in certain ways as well as “victimized” by what happened to them as young children (i.e., psychodynamic theory). Mischel acknowledged that the early personality researchers would mention the role of situations but would still revert to the idea that personality is completely internal. Rather, Mischel advocated that we have done a poor job of taxonomizing environmental influences on human behavior and thus need a better way of thinking about human action.

Mischel (1977) defined a strong situation as one that provides behavioral norms, explicit informational cues and incentives to individuals in order to fulfill task requirements. One would witness low variation among individuals within a situation. For a situation to be powerful, above and beyond individual differences, one would notice uniform encoding of events, uniform expectancies, incentives for performance, and individuals must possess the requisite skills to perform the given task. Examples of strong situations may include funerals, trauma-induced responses, rigid instructions, etc. A weak situation, on the other hand, allows individuals greater discretion in how to operate within the confines of the situation and offer less direction and more ambiguity in how situational cues are interpreted. One would witness high variation among individuals within a given situation. Weak contexts have the greatest potential to reveal one’s personality traits, assuming the traits are relevant to the given situation.
Upon viewing the differences in structure between error management training and error avoidant training, one can make competing arguments. One, because error management training is ill-structured in nature and error avoidant training provides strict guidelines to follow, one could surmise that error management training is a weak situation allowing for individual differences to influence its effectiveness whereas error avoidant training would exemplify a strong situation. On the other hand, the error management training framing statements (e.g., “You made a mistake! Great!”) might be an example of a strong situation such that personality might be rendered irrelevant in influencing thought and behavior. The purpose of the framing statements are intended to attenuate anxiety and frustration levels associated with mistakes and errors. This forces the trainee to attend to the task and avoid personal rumination. The absence of framing statements in error avoidant training might allow greater discretion on behalf of individual thought and behavior to influence performance effects. Let us examine previous research before definitively deciding which to hypothesize.

Gully et al. (2002), for instance, examined the interactions of training type with conscientiousness and openness to experience on three different outcomes: declarative knowledge, task performance, and self-efficacy. Results supported the openness to experience hypothesis but only supported conscientiousness as a moderator of training type effects on self-efficacy. The operationalization of the manipulations was similar to previous studies but deviated from traditional error management training operations because the framing statements were absent in Gully et al. (2002). The manipulation description encouraged the positive aspects of errors and exploring the causes of those errors, but the absence of persistent reminders may have weakened the strength of the manipulation. This idea is expounded upon in Heimbeck et al. (2003).
Heimbeck et al. (2003) expected to see an attribute-treatment interaction such that individuals with high trait learning goal orientations would benefit more from error management training than error avoidant training. They also expected individuals with high trait performance goal orientations to benefit more from error avoidant training than error management training. However, the authors only found a significant effect between performance goal orientations and error avoidant training. To explain the null results for error management training, Heimbeck et al. (2003) reasoned that error management training provided strong cues to regulate participants’ thoughts and behavior in a training session. The exploratory condition might provide better evidence of a weak situation and determine whether the framing statements cue individual behavior and render personality irrelevant in the error management training domain. We would expect error management training with framing statements to be more immune to personality influences than error management training without framing statements. Therefore, we suggest that researchers should switch focus to the interactionist perspective and embrace Mischel’s (1968) concept of situation strength.

To address this issue, I suggest that in order to test the assumption that personality moderates the effects of training type on performance, one must separate the training types into three conditions. To determine the situation strength of error management training, I formed an error management training condition (exploration plus framing statements), exploratory learning condition (no framing statements), and the error avoidant training condition. Thus, we would expect personality to have the largest effect on performance within the exploratory learning condition, followed by the error avoidant training condition. Personality should not affect performance in the error management training condition because there are more environmental cues present than in either the exploratory learning or error avoidant training condition to
influence participant behavior. Let us now look at two personality constructs that have not been advanced yet in the error management training literature and also may bear relevance within this context.

**Perfectionism.** An initial search into the perfectionism literature conjured a wealth of knowledge from the clinical side of psychology ranging from studies devoted to eating disorders (Downey & Chang, 2007; Castro-Fornieles et al., 2007; Macedo et al., 2007), clinical depression (Harris, Pepper & Maack, 2008; Clara, Cox & Enns, 2007) and Obsessive-Compulsive Disorder (Knutt, 2007; Radomsky et al., 2007). In essence, a majority of researchers has investigated the construct from a psychopathological perspective, treating it more as a personality disorder than a personality construct. Other researchers have questioned the one-sidedness of this conceptualization. Recently, researchers have been arguing that there is an adaptive side to perfectionism (Flett & Hewitt, 2002; Frost, Marten, Lahart, & Rosenblate, 1990).

Indeed, there have been a few instances in recent years where perfectionism has been addressed within the I/O literature (Simmons, Watson, & O’Leary, 2007; Trumpeter, Watson, & O’Leary, 2006). Arguably, one needs a certain level of perfectionistic strivings to attain one’s goals. Frost, Marten, Lahart and Rosenblate (1990) developed the Multidimensional Perfectionism Scale (MPS) to assess two factors of perfectionism which they term dysfunctional and healthy (later researchers have used the term unhealthy instead of dysfunctional as well as maladaptive/adaptive). Flett and Hewitt (2002) defined adaptive perfectionism as self-oriented and reflecting high personal standards whereas maladaptive perfectionism evokes a high concern over mistakes and is socially-oriented in nature. Hamachek (1978) described these dimensions in more detail. Hamachek advocated the terms normal and neurotic to distinguish the two types of perfectionism in order to highlight the positive side of the construct. Normal perfectionists
strive for reasonable and realistic standards that enhance the sense of self. Neurotic perfectionists have excessively high standards and are governed by the thoughts and perceptions of others (e.g., performance evaluation).

Stoeber and Otto (2006) addressed two approaches to categorizing perfectionism. One is a dimensional approach, which portrays perfectionistic strivings and perfectionistic concerns, the former being positive and the latter being negative. The former is also self-oriented with high personal standards whereas the latter is concerned with mistakes, doubts about actions, and socially-prescribed. The other approach is uni-dimensional (or group-based). It classifies perfectionism along the lines of healthy (high striving, low concern) and unhealthy (high striving, high concern). Flett and Hewitt (2002) advocated the use of the dimensional approach because there are certain degrees of perfectionism that individuals possess as opposed to the either/or aspect of the group-based approach. Trumpeter, Watson, and O’Leary’s (2006) findings also lend support to this interpretation. On the basis of prior research, I used the dimensional approach.

Individuals with a high focus on errors/mistakes would likely show a decrement in performance, specifically in an error training format. The issue is whether or not the error management training manipulation is strong enough to counter the irrational fear of failure on behalf of the maladaptive perfectionists. This issue is one which is addressed in this study. Based on the above discussion in regards to perfectionism, I offer the following hypothesis.

H5a: Adaptive perfectionism will be positively related to transfer performance.

H5b: Maladaptive perfectionism will be negatively related to transfer performance.

H5c: Adaptive perfectionism will interact with training type effects on transfer performance such that adaptive perfectionism will be positively related to transfer performance in the
exploratory learning and error avoidant training conditions but unrelated in the error management training condition.

$H_{5d}$: Maladaptive perfectionism will interact with training type effects on transfer performance such that maladaptive perfectionism will be negatively related to transfer performance in the exploratory learning and error avoidant training conditions but unrelated in the error management training condition.

**Optimism.** Scheier and Carver (1985) defined optimism as a general belief that one holds about the world, that people will generally experience good vs. bad outcomes. Optimistic individuals tend to stay positive, even when negative occurrences beset them or others, and believe that the world has a natural balance of positive and negative. Optimism is a unidimensional construct, with pessimism as its polar opposite. Carver and Scheier (2003) differentiated optimism from constructs such as personal control and efficacy in that the former is a much broader construct. Where the constructs converge is in relation to effort. Optimists do not sit idly and wait for outcomes to occur. On the contrary, optimists still believe that effort remains an integral part of valued outcomes, just as the self-efficacy construct would posit. Indeed, in a study discussing attributions of failure and future corrective actions, salespeople who exhibited more optimism put forth more effort on the next sales call than salespeople who were pessimistic (Dixon & Schertzer, 2005).

Optimism differs from self-efficacy, though similar in origin (expectancy perspective), in two ways. One, Scheier and Carver (1992) attested that Bandura (1977) viewed the self-efficacy construct as a central determinant of behavior whereas optimism accounts for other sources of variance (e.g., religious faith, available assistance from other people, etc.). Secondly, self-efficacy is more situation-specific whereas optimism entails a more generalized expectancy.
As previously mentioned, Carver and Scheier (2003) posited that optimism is rooted in the expectancy-value model, as opposed to a personality point of view. The expectancy-value model suggests that people are motivated to put forth effort towards tasks or activities for which they believe that effort will lead to some valued outcome. Optimists tend to have higher expectancies than pessimists in relation to goal pursuit.

Optimism has been shown to predict many valuable work-related outcomes. For instance, optimism predicted job performance (Arakawa & Greenberg, 2007; Youssef & Luthans, 2007), job satisfaction, and work happiness (Youssef & Luthans, 2007). Optimism also moderated the training intervention-performance relationship in a vigilance training exercise (Szalma, Hancock, Dember, & Warm, 2006). Optimists also tended to report less job stress than pessimists (Tuten & Neidermeyer, 2004), and they coped with failures better than pessimists by employing more problem-focused strategies as opposed to emotion-focused strategies (Strutton & Lumpkin, 1993). In an error management training intervention, trainees must be able to persist on task-related activities in order to enhance learning and growth activities because error management in-training performance tends to be stressful and lower than other, more traditional forms of training. Along the lines of previous reasoning, I do not expect optimists to be more adaptive than pessimists in the error management training condition (due to the strong situational nature of the intervention). Assuming optimists have more adaptive mechanisms to counter personal failures, mainly the control of negative emotions and general stress, I expect that optimists, in an exploratory learning and error avoidant training condition, will perform better than pessimists.

H_{6a}: Optimism will be positively related to transfer performance.
H_{6b}: Optimism will interact with training type effects on transfer performance such that optimism will be positively related to transfer performance in the exploratory learning and error avoidant training conditions but unrelated in the error management training condition.

A full conceptual model of the variables and relationships within this study can be seen in Figure 1.

*Figure 1 Conceptual Model*
Method

Participants

I sampled from a student population in a mid-size, Midwestern university. To determine the required sample size in order to ensure sufficient power to detect an effect, I conducted a power analysis. To accomplish this I set alpha at .05, the effect size at .20 for a small effect (Cohen, 1988) and expected power to be at least .80. From this information, I determined that the required sample size for this study was 193. Therefore, I set out to sample approximately 200 participants for this study. Usable data after data cleaning accounted for slightly less than what the power analysis called for (n = 181). Because undergraduate women at this university outnumbered the men by a ratio of approximately 2:1, and specifically, the females in the psychology subject pool on average account for approximately 65-70% of the total participants, I intended to ensure the same proportion of participants by gender in each treatment condition. There were 109 females in the sample (60%) and these percentages were approximately equal across conditions.

Experimental Design

Each participant received one of the following three treatment conditions: error-management training, exploratory learning, and error-avoidant training. All participants performed four training trials followed by one transfer trial. In addition, I examined the effects of training type on three mediators (meta-cognition, emotion control, and cognitive appraisals, respectively). Further, I examined perfectionism and optimism as moderators of training type effects on performance.

Task Description
Participants performed a computerized class scheduling task, similar to the one used in Mangos and Steele-Johnson (2001). This task was moderately difficult to perform given that cognitive ability predicts performance on this task (see results). Although participants were students, and presumably have had some task experience with scheduling classes, participants nonetheless were at least unfamiliar with the task interface (MS-DOS) and thus required some learning. In this task, individuals formed class schedules for mock students according to a pre-specified set of rules (see Appendix A). Participants chose courses from a database and arranged them in such a way as defined by the rules. The task window was divided into four quadrants. The upper left quadrant is called the Course Schedule Window. In this quadrant, participants could view different courses and different sections within each course by using the PAGE UP and PAGE DOWN functions and the UP and DOWN arrows, respectively. The upper right window is called the Planning Window. Here, participants could view the entered courses from the course schedule window. In the lower right quadrant, participants would look at the Review Window, where completed schedules could be reviewed. Finally, in the lower left quadrant, participants would see the Information Window. In this quadrant, the number of errors and number of completed schedules are displayed in this portion of the screen. Also, when a rule is violated, the exact rule would be highlighted in red in this quadrant. Participants could recall any rule by hitting the F1-F7 keys.

I created two versions of the task: a training version and a transfer version. Transfer tasks are often structurally different and more complex than the training version. Past studies have asked participants to perform functions not learned in training, such as re-creating slides in PowerPoint during training and creating colored text boxes in the transfer phase (Keith & Frese, 2005). I distinguished the types of task based on its complexity by adding more rules that would
constrain participant performance. This increase in task complexity likely would challenge participants to apply the strategies formed during the training task. The training task consisted of six rules, whereas the transfer task consisted of seven rules (Mangos & Steele-Johnson, 2001).

**Manipulation Description**

There were three types of training, the first of which was called *error-management training* (EMT). This condition required at least two important features. One, experimenters allowed and encouraged the exploration of the task and did not aid participants with solutions or hints that would guide behavior. The experimenter limited help and focused upon error-encouragement, and experimenters intervened only for non-task related problems (e.g., computer malfunction). Two, the EMT condition contained positive framing statements. The framing statements were meant to encourage and enable both the occurrence and management of errors. The framing statements have been designed to neutralize the negative emotional arousal and thoughts that often accompany errors. Experimenters attached the framing statements to the computers so as to make them visually salient to participants, and the statements were repeated when participants inquired about the task for clarification. The set of heuristics used in this study (Heimbeck et al., 2003) are shown in Table 1. Also, the text of the manipulation description can be seen in Appendix B.

The second condition, called *exploratory learning* (EXP), resembled the EMT condition but did not include framing statements. Participants were encouraged to explore the task but were given no instructions in regard to the valence of errors (see Appendix B).

The third experimental condition was called *error-avoidant training* (EAT). I conceptualized this condition as following a step-by-step protocol. If an error was made, the experimenter would encourage the individual to abide by the protocol or follow the directions
given to them. This condition discouraged participants from exploring the causes of mistakes (see Appendix B).

Table 1

*EMT Framing Statements*

1. Errors are a natural part of the learning process!
2. Errors inform what you are still able to learn!
3. There is always a way to leave an error situation!
4. The more errors you make, the more you learn!

To test the effectiveness of the manipulation induction and its effects on performance, I conducted two pilot studies (see Appendix C).

For a simplified, organized view of the manipulation, see Table 2.

Table 2

*Manipulation Components by Condition*

<table>
<thead>
<tr>
<th></th>
<th>Error Valence</th>
<th>Learning Method</th>
<th>Framing Heuristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMT</td>
<td>“…when you encounter an error, do not feel stressed or dwell on it…errors are natural part of learning…”</td>
<td>“…feel free to explore the task and discover why you made that error…understanding why will help.”</td>
<td>Ex. “The more errors you make, the more you learn.”</td>
</tr>
<tr>
<td>EAT</td>
<td>“…you want to avoid mistakes at all costs. Errors force you to work inefficiently and decrease your performance.”</td>
<td>“…make sure you follow instructions given to you…rules and guides will help learning the task…”</td>
<td>NONE</td>
</tr>
<tr>
<td>EXP</td>
<td>NONE</td>
<td>“…feel free to explore the task…discover as much as you can…”</td>
<td>NONE</td>
</tr>
</tbody>
</table>
Measures

Manipulation check. To ensure that the induction had the intended psychological effect, I conducted a check of the manipulation. The check assessed the strength of the manipulation by asking participants the extent to which they viewed errors either positively or negatively. EMT participants should show a higher propensity to endorse positive views of errors. For this study, I followed the same procedure used by Keith and Frese (2005) and used the Error Orientation Questionnaire (EOQ) developed by Rybowiak, Garst, Frese, and Batinic (1999). The only difference between Keith and Frese (2005) and this study is that I adapted the items to reflect task-specificity. The EOQ consisted of eight subscales, of which only two were relevant for this study. These two subscales were administered for the manipulation check (i.e., Error Strain, Learning from Errors). Error Strain is a 7-point Likert-type scale consisting of five items asking participants the extent to which the statements apply to them (1 = strongly disagree, 7 = strongly agree). Error Strain items reflect a negative endorsement of errors. Learning from Errors is also a 7-point Likert-type scale but consists of only four items. The Learning from Errors subscale signifies a positive view of errors. The full manipulation check can be seen in Appendix D. Reliability alphas for the English versions of the two subscales are .89 and .79 (Learning from Errors and Error Strain, respectively; Rybowiak et al., 1999).

To assess the manipulation more directly, I also added three additional items that can be seen in Appendix D. My intention for using these additional items was to achieve a more direct observation of the manipulation within the confines of the individual just in case the original manipulation check was too indirect. Factor analyses and reliabilities of revised scale can be found in the results section.
Demographics. I obtained demographic information (e.g., gender, race, class rank, etc.) in order to rule out alternative explanations for the results (see Appendix E).

Perfectionism. I assessed perfectionism using the Almost-Perfect Scale-Revised (Slaney, Rice, Mobley, Trippi, & Ashby, 2001). This scale (see Appendix F) conceptualized perfectionism as a three-dimensional construct, consisting of high standards, order, and discrepancy. Slaney et al. (2001) reported internal consistency reliabilities of .85, .82, and .91, respectively. High standards and order lack negative connotations and are considered to represent positive striving in one’s life. Discrepancy, on the other hand, is deemed to represent maladaptive evaluations. The scale is a 23-item measure and participants respond using a 7-point Likert-type scale (1 = strongly disagree, 7 = strongly agree). There are twelve items measuring discrepancy, seven items measuring high standards, and four items measuring order. Subscale scores are computed by summing item responses and none of the items are reverse-scored.

Researchers have combined high standards and order to reflect adaptive perfectionism whereas discrepancy has been suggested to reflect maladaptive perfectionism. In pilot research, using a sample drawn from the same population as the current study, I ran an exploratory factor analysis (EFA) to see how many factors were reflected in the data and then constrained a subsequent analysis to the hypothesized factor structure. For the EFA, I accepted eigenvalues greater than 1.0 and determined that the scale had four factors. However, a closer examination of the factor loadings revealed weak loadings of items on the fourth factor and suggested the acceptance of a three factor structure. I ran two subsequent analyses to test whether the data fit well to a 2- (as conceptualized in the literature) or 3- (as empirical evidence would suggest) factor structure. According to the results, either structure would suffice. Therefore, I accepted
the more parsimonious solution that agreed with a common conceptualization of the perfectionism construct (i.e., two factor structure). The two-factor structure explained 50% of the observed variance in the perfectionism construct. The eigenvalues for a one, two, and three factor structure suggested that a two-factor structure is reasonably justified (6.5, 5.0, and 1.95, respectively). Plus, adding a third factor would increase the variance accounted for by only 8%. For the purposes of this study, I accepted perfectionism as having a two-factor structure (for detailed tables and scree plots, please see the rest of Appendix F for the appropriate information).

Reliability alphas for this study were .92 for maladaptive perfectionism and .85 for adaptive perfectionism.

**Optimism.** I assessed optimism using the Life Orientation Test-Revised (LOT-R) (Scheier, Carver, & Bridges, 1994). This scale (see Appendix G) is uni-dimensional, with optimism on the positive end and pessimism reflecting the lower end. The LOT-R is a 10-item measure on a 7-point Likert-type scale (1 = strongly disagree, 7 = strongly agree). Internal consistency reliability for the LOT-R is .78, and test-retest reliability coefficients are .68 (4 months), .60 (12 months), .56 (24 months), and .79 (28 months) (Scheier, Carver, & Bridges, 1994). For this study, I achieved a reliability alpha of .76.

**Meta-cognition.** I used an adaptation of the Cartwright-Hatton and Wells’ (1997) Meta-Cognitions Scale for Worry and Intrusive Thoughts. Specifically, I used the Cognitive Self-Consciousness subscale of the measure. Cartwright-Hatton and Wells’ (1997) scale was intended to measure meta-cognitions as a disposition. I revised the items to reflect the state aspects of meta-cognitions by altering the items from present to past tense and qualifying each statement with a clause that made the task salient (see Appendix H). The measure is a 7-item, 7-point Likert-type scale (1 = strongly disagree, 7 = strongly agree). Cartwright-Hatton and Wells
(1997) reported internal consistency reliability for their original measure of Cognitive Self-Consciousness as .72. Reliability alphas for the revised measure in this study were .86 at Time 1 and .88 at Time 2 (please see Appendix H for revisions to the scale).

**Emotion control.** I used the measure of emotion control developed by Keith and Frese (2005). This measure (see Appendix I) contains eight items, and participants respond on a 5-point Likert-type scale (1 = false, 5 = true). Internal consistency reliability for this measure was .82 for the initial pilot test and .80 for the Keith and Frese (2005) sample. For this study, the reliability alphas were .76 at Time 1 and .85 at Time 2.

**Cognitive appraisals.** I used the cognitive appraisal measure from McGregor and Elliot (2002). This measure (see Appendix J) consists of five items assessing challenge and four items assessing threat. Participants respond on a 7-point Likert-type scale (1 = strongly disagree, 7 = strongly agree). Cronbach alphas for the measure are .85 and .80 (Threat and Challenge, respectively) (McGregor & Elliot, 2002). For challenge appraisals, reliability alphas for this study were .82 at Time 1 and .85 at Time 2. For threat appraisals, reliability alphas were lower than the alpha reported by McGregor and Elliot (2002). At Time 1, alpha was .61 and at Time 2 alpha was .64.

**Performance.** I operationalized task performance as the number of schedules completed within each trial session, which reflected a quantity measure of performance. I also measured performance as number of errors within each trial session, which reflected a quality measure of performance.

**Additional Exploratory Measures**

I administered the following measures to enable exploratory analyses in preparation for future research. These measures are not reported in the current study.
Self-set goals and goal commitment. I measured self-set goals (see Appendix K) with a one item measure asking participants about how many schedules they think they can make in a subsequent trial. I measured goal commitment as a follow-up to the self-set goal assessment using the 7-item self-report measure designed by Hollenbeck, Klein, O’Leary, and Wright (1989). This measure is a 7-point Likert-type scale (1 = strongly disagree, 7 = strongly agree). Internal consistency reliability for this measure based on the Hollenbeck et al. (1989) sample is .80. Reliability alphas for this particular study were .69 at Time 1 and .75 at Time 2.

Subjective task complexity. I used a measure of subjective task complexity taken from a study by Steele-Johnson, Beauregard, Hoover, and Schmidt (2000) (see Appendix L). Participants responded to 7 items on a 7-point Likert-type scale (1 = strongly disagree, 7 = strongly agree). Items are in the form of a question that pertained to the relative difficulty of the task. Steele-Johnson et al. (2000) attained internal consistency reliabilities of .77, .76, .76, and .71 in four successive trial blocks. For this particular study, the reliability alphas were .63 at Time 1 and .69 at Time 2.

Task satisfaction. I also used a measure of task satisfaction (Appendix M) taken from Steele-Johnson et al. (2000) which contained two items. Participants responded to this measure of task satisfaction on a 7-point Likert-type scale (1 = strongly disagree, 7 = strongly agree). Internal consistencies for four successive trial blocks, according to Steele-Johnson et al. (2000) were .71, .56, .72, and .72. For this particular study, reliability alphas were .66 at Time 1 and .71 at Time 2.

Self-efficacy. I used the measure developed by Riggs, et al. (1994) (see Appendix N). The Riggs et al. (1994) measure of self-efficacy is 9 items, and participants respond on a 7-point Likert-type scale (1 = strongly disagree, 7 = strongly agree). Internal consistency reliability,
according to Riggs et al. (1994), is .86. According to this study, reliability alphas were .78 at Time 1 and .72 at Time 2.

Procedure

At the beginning of the experiment, researchers provided informed consent forms to participants (see Appendix O) and were seated individually at a computer kiosk. Each session had six participants each. Once there, participants completed the following personality measures: perfectionism, optimism, and error orientation. At the completion of the personality questionnaires, participants listened to audio-recorded task instructions (see Appendix P). These instructions served the purpose of familiarizing the participants with the basic functions of the task. To go along with the audio instructions, experimenters handed out three sets of sheets to participants. One was a diagram of the keyboard with relevant keys highlighted. Another was a screen shot of the course planning window (see Appendix P). Finally, the third was a packet of the typed instructions that participants listened to verbatim. The intention of the audio recording was to ensure equal exposure to the instructions among all participants. Once the audio completed, participants turned the instructions back in to the experimenter and did not have access to the instructions later on in the experiment.

After task instructions were completed, participants read the task rules on the computer. Once every participant had read the task rules, participants received one of the three manipulation inductions. The experimenter read the manipulation induction out loud once in its entirety, and then again line by line so that participants could copy down the induction on a sheet of paper provided to them. This was to ensure that participants were encoding the manipulation information properly.
Participants then received a feedback screen prompt which was tailored to each manipulation condition (see Appendix Q). After this exposure, participants began the first training trial. Each trial is 10 minutes in duration and participants create as many schedules as they can in any given trial. At the end of the first trial, participants filled out the first administration of the manipulation check. Because the manipulation check was task-specific, I required participants to complete one trial of the task before completing the measure. After this, the experimenter read aloud the manipulation paragraph once again to strengthen the effects of the induction (participants did not copy down the paragraph this time). From here, participants completed three more training trials of the task, with each trial preceded by the feedback screen. After completion of the fourth trial, participants completed the manipulation check once again to assess the strength of the manipulation over time. Following the second administration of the manipulation check measure, participants filled out the following questionnaires: cognitive appraisals, meta-cognitions, emotion control, subjective task complexity, and task satisfaction. Prior to beginning the next trial, participants completed three future-oriented measures: self-efficacy, self-set goals, and goal commitment.

Once participants completed the aforementioned measures, they received instructions in regard to the transfer trial. Experimenters informed the participants that the new trial (i.e., transfer trial) is a new version of the task that assessed competence learned in previous trials and presented a more difficult challenge for participants than experienced in previous trials. Operationally, there is one more rule added to the task to increase the complexity of the task, thus creating more difficulty and offering a better test of strategy development between the experimental conditions. Participants re-read the task rules (including the new one) and then performed the 10-minute transfer task two times. Following the second transfer trial,
participants completed the next round of measures, which were in the same order as previously mentioned. In addition to the previous measures, participants completed a goal orientation measure, demographics measure and self-reported ACT/SAT form for exploratory analyses. Finally, the experimenter debriefed the participants (see Appendix R) and thanked them for their participation. For a brief outline of the procedure, see Table 3.

Table 3

*Experimental Procedure Outline*

1. Perfectionism and Optimism
2. Task instructions
3. Task rules
4. Manipulation
5. 1st Training trial
6. Manipulation check
7. Repeat manipulation
8. 3 Training trials
9. Manipulation check and mediation measures
10. Transfer screen and transfer task rules
11. 2 Transfer trials
12. Mediation Measures
Results

Sample Characteristics

Sample Size. Two hundred eighteen participants were available for analysis. Of these, 37 (17% of total sample) were removed from the sample, resulting in 181 participants. Sixteen participants (43% of those removed) were removed due to patterned responding. Seven participants (19%) were outliers and caused unnecessary skewness to the data (e.g., accruing 1200 errors in one trial). Six participants (16%) had missing data. Four participants (11%) failed to follow task instructions during the experiment and thus were removed from the sample. Two participants (5%) were monitored and assessed as having collaborated together when the experimenter was not aware. Therefore, we deleted these two participants due to potential confound problems in their data. One participant (3%) had language problems (i.e., English was not his/her first language) so that person’s data was not used. Finally, there was an experimenter error in protocol in which information was given to the participant that should not have been, so I deleted this person’s data (3%) due to possible confound problems.

Participants were removed approximately equally across conditions. Twelve participants were removed from the EMT condition, 15 from the EAT condition, and 10 from the EXP condition.

Means, standard deviations and intercorrelations of all variables can be found in Table 4.

Initial group differences. The EMT and EAT conditions were approximately equal in size (n = 65, n = 64, respectively) whereas the EXP condition had slightly fewer participants (n = 52), although not at a significant level, $\chi^2(2) = 1.74, p = .42$. The gender distribution was as expected; there were more females than males (female n = 109, male n = 72) which was consistent with the proportion of females in the subject pool. Gender was distributed similarly
Table 4

Means, Standard Deviations, Reliability Alphas and Correlations of Study Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>SCH</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ERR</td>
<td>.04</td>
<td>.19**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>EMOTC</td>
<td>.05</td>
<td>.05</td>
<td>-.13^</td>
<td>(.76)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>METAC</td>
<td>-.02</td>
<td>-.05</td>
<td>.10</td>
<td>.21**</td>
<td>(.86)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ADPER</td>
<td>-.02</td>
<td>.01</td>
<td>.05</td>
<td>.17*</td>
<td>.19**</td>
<td>(.85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>MALPER</td>
<td>-.09</td>
<td>-.00</td>
<td>.01</td>
<td>-.25**</td>
<td>.14*</td>
<td>.05</td>
<td>(.92)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>OPTM</td>
<td>.12</td>
<td>.06</td>
<td>.12</td>
<td>.26**</td>
<td>.05</td>
<td>.24**</td>
<td>-.54**</td>
<td>(.61)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>THREAT</td>
<td>-.16*</td>
<td>-.22**</td>
<td>.00</td>
<td>-.38**</td>
<td>.03</td>
<td>-.07</td>
<td>.40**</td>
<td>-.26**</td>
<td>(.61)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>CHALL</td>
<td>-.09</td>
<td>-.09</td>
<td>-.01</td>
<td>.38**</td>
<td>.42**</td>
<td>.27**</td>
<td>.01</td>
<td>.22**</td>
<td>.17*</td>
<td>(.82)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ES</td>
<td>-.12</td>
<td>-.16*</td>
<td>.08</td>
<td>-.22**</td>
<td>.13^</td>
<td>.05</td>
<td>.39**</td>
<td>-.22**</td>
<td>.49**</td>
<td>.15*</td>
<td>(.82)</td>
</tr>
<tr>
<td>12</td>
<td>LE</td>
<td>.11</td>
<td>.03</td>
<td>-.02</td>
<td>.26**</td>
<td>.12</td>
<td>.24**</td>
<td>-.16*</td>
<td>.25**</td>
<td>-.17*</td>
<td>.29**</td>
<td>-.21**</td>
</tr>
</tbody>
</table>

Mean                  1.93 | 4.43 | 1.59 | 4.25 | 4.78 | 5.68 | 3.62 | 4.69 | 2.70 | 4.68 | 3.83 | 4.99
Standard Dev.         0.80 | 2.05 | 1.68 | 0.71 | 1.34 | 0.87 | 1.36 | 1.14 | 1.14 | 1.31 | 1.52 | 1.22

Note:   
N = 181; ^p < .10, *p < .05, **p < .01, alphas in parentheses
Variables 2 and 3 are average performances for the two transfer trials
Variables 4, 5, 9, 10, 11, 12 are measures at Time 1
TT = Training Type; SCH = Average number of schedules produced during transfer trials; ERR = Average number of errors committed during transfer trials; EMOTC = Emotion Control; METAC = Metacognition; ADPER = Adaptive Perfectionism; MALPER = Maladaptive Perfectionism; OPTM = Optimism; THREAT = Threat Appraisal; CHALL = Challenge Appraisal; ES = Error Strain; LE = Learning from Errors

across conditions, $\chi^2(2) = 1.00$, p = .61. In the EMT condition, there were 39 female and 26 male participants. Similarly, there were 36 female and 28 male participants in the EAT condition.

Finally, in the EXP condition, there were 34 female and 18 male participants.

In addition to assessing the gender distribution in each condition, I also analyzed initial group differences based on cognitive ability to ensure that random assignment was effective in distributing cognitive ability similarly across conditions. Indeed, I found this to be the case.
EMT ($M = 22.53$, $SD = 4.69$), EAT ($M = 22.33$, $SD = 3.80$), and EXP ($M = 22.42$, $SD = 3.76$) participants did not differ based on self-reported ACT scores, $F(2, 139) = 0.031$, $p = .97$. One caveat with this analysis is that 39 people failed to report their ACT scores (possibly due to forgetting or not having taken the test). Thus, the sample size was reduced to 142 for this analysis. There were 16 participants in both the EMT and EAT condition who failed to report their ACT scores, whereas only seven participants failed to report their ACT score. This resulted in similar composition between each condition for the ACT analysis (EMT $n = 49$, EAT $n = 48$, EXP $n = 45$). Cognitive ability also predicted transfer performance (schedules), $F(1, 140) = 6.69$, $p < .01$.

**Manipulation Check**

*Psychometrics.* I used two subscales from the Error Orientation Questionnaire (Rybiowak et al., 1999) to assess the manipulation and adapted the items to reflect task-specificity. Specifically, there were four items measuring Learning from Errors and five items measuring Error Strain. Additionally, on the basis of committee recommendation, I had added three items that were more sensitive to the manipulation. I assessed the factor structure of all twelve items and expected to observe two factors with the three additional items having high loadings on the Learning from Errors factor and low loadings on the Error Strain factor.

First, I ran an unrotated principal components analysis to assess the number of factors that were explained by the data. Any factors with eigenvalues below 1.0 were not considered. Examination of the eigenvalues suggested that three factors were possible. Factor 1 had an eigenvalue of 4.149 and explained 35% of the variance. Factor 2’s eigenvalue was 2.724 and explained 23% of the variance. However, factor 3’s eigenvalue was 1.059 and only explained
9% of the variance. Thus, based on the combined proportion of variance explained by factors 1 and 2 (57%) and the scree plot, I concluded that a two-factor structure was defensible.

However, before coming to a final conclusion, I constrained the data to 1, 2 and 3 factor solutions in an unrotated factor analysis to confirm that the factor loadings of the items were as hypothesized. I considered any factor loading of .30 or larger to be acceptable. In a one factor solution, the hypothesized Learning from Errors items produced factor loadings that were all positive and larger than .30 whereas the Error Strain items produced factor loadings that were all negative and smaller than -.30. Next, I assessed the two-factor solution. Forcing a two-factor solution revealed high loadings for the seven Learning from Errors items on one factor and high loadings for the five Error Strain items on the second factor. Moreover, the items loading on one factor had low loadings on the second factor.

To ensure that a three factor solution did not fit the data better than a two factor solution, I assessed the factor loadings of the three factor solution. Again, I witnessed a similar pattern of results as explained by the two factor solution. The only difference was revealed for Item 10, which had high loadings on both Factor 1 and Factor 3. This item was one of the additional items designed to assess the manipulation. This was the only item with a high loading (.63) on the third factor. The rest were below the minimum cut-off (< .30).

Given that there was a small, negative correlation between the factors, I elected to employ an orthogonal rotation on the factor structure using the varimax procedure. This procedure ensures that the factors are uncorrelated. I ran the varimax rotation factor analysis with 1, 2 and 3 factor solutions. I noticed no differences between the unrotated and rotated factor solutions for either 1 or 2 factors. For the 3 factor solution, Item 10 was related more strongly to the third factor and its loading on Factor 1 was smaller than .30. The other two
additional items each had high cross loadings between Factors 1 and 3. However, having a one item factor does not make much sense psychometrically, given that one item measures tend to exhibit low reliability. Deleting the other two items would lose predictive power, because the items loaded onto Factor 1 quite well. According to the proportion of explained variance and the scree plot, I felt justified in treating this measure as having a two-factor solution. For the factor loadings of the two-factor varimax solution, see Table 5.

I computed four internal consistency reliabilities for the two scales at Time 1 and Time 2. Learning from Errors at Time 1 ($\alpha = .85$) and Time 2 ($\alpha = .84$) and the Error Strain scales at Time 1 ($\alpha = .82$) and Time 2 ($\alpha = .84$) all displayed adequate reliabilities. Therefore, I could proceed reasonably to assessing the efficacy of the manipulation using these measures.

*Test of Transfer Task Operationalization.* By increasing the transfer trial by one rule, I expected the task to be more complex and difficult. Thus, I would expect performance to decrease from Trials 4 ($M = 6.54, SD = 3.11$) to 5 ($M = 4.27, SD = 2.30$). Indeed, performance significantly decreased, $t(180) = 11.41, p < .0001$.

*Manipulation check by condition.* To assess the effectiveness of the manipulation, I ran a one-way ANOVA on the two subscales at Time 1 and Time 2 with condition as the independent variable. Of the four omnibus tests, only the Error Strain measure at Time 1 was significant, $F(2, 178) = 4.46, p = .01$. Learning from Errors at Time 1, $F(2, 178) = 1.81, p = .17$, Error Strain at Time 2, $F(2, 178) = 2.01, p = .14$, and Learning from Errors at Time 2, $F(2, 178) = 1.87, p = .16$, were all non-significant. The means by condition were all in the expected direction (see Table 6).
Post-hoc comparisons suggested that the two conditions of particular interest to the study (EAT and EMT) showed significant differences. There were two different contrasts used in the post-hoc tests. The first one compared EAT with EMT and EXP as one comparison group. It was thought that the latter two groups would share more in common than the EAT group. The second contrast compared EAT with EMT. Although these two contrasts are not orthogonal, these contrasts reflect the comparisons most central to my study.

For the Error Strain at Time 1 measure, EAT participants experienced significantly more strain than the EMT and EXP participants combined, \( t(178) = -2.60, p < .01 \) (for means and standard deviations, see Table 2). For the second contrast, as expected, EAT participants experienced more strain than EMT participants, \( t(178) = 2.99, p < .01 \).

For the Error Strain at Time 2 measure, EAT participants again experienced more strain when faced with errors than EMT and EXP combined, \( t(178) = -2.00, p = .047 \). However, when comparing EAT and EMT participants, they did not differ significantly, \( t(278) = 1.68, p = .09 \).

For the Learning from Errors at Time 1 measure, although EMT and EXP participants combined felt one could view errors more positively than EAT participants, they did not believe so at a significant level, \( t(178) = 1.86, p = .06 \). Again, EMT participants tended to view errors more positively than EAT participants, but not at a significant level, \( t(178) = -1.79, p = .07 \).

Finally, for the Learning from Errors at Time 2 measure, EMT and EXP participants did not differ from EAT participants in their assessment of errors, \( t(178) = 0.41, p = .68 \). Further, EMT participants did not differ from EAT participants in how positively they viewed errors, \( t(178) = -1.35, p = .18 \).
Table 5

*Factor Analysis with Varimax Orthogonal Rotation (2-factor solution)*

<table>
<thead>
<tr>
<th>Items</th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>I found it stressful when I made an error on the task.</td>
<td></td>
</tr>
<tr>
<td>I was often afraid of making mistakes on the task.</td>
<td></td>
</tr>
<tr>
<td>I felt embarrassed when I made an error on the task.</td>
<td></td>
</tr>
<tr>
<td>If I made a mistake, I “lost my cool” and became angry while working on the task.</td>
<td></td>
</tr>
<tr>
<td>While working on the task, I was concerned that I did something wrong.</td>
<td></td>
</tr>
<tr>
<td>Mistakes assisted me to improve my performance on the task.</td>
<td></td>
</tr>
<tr>
<td>Mistakes provided useful information for me while working on the task.</td>
<td></td>
</tr>
<tr>
<td>My mistakes helped me to improve my performance on the task.</td>
<td></td>
</tr>
<tr>
<td>My mistakes had helped me to learn the task better.</td>
<td></td>
</tr>
<tr>
<td>While working on the task, I was able to maintain task focus.*</td>
<td></td>
</tr>
<tr>
<td>I felt positive after making a mistake on the task.*</td>
<td></td>
</tr>
<tr>
<td>While working on the task, I was encouraged to explore the task.*</td>
<td></td>
</tr>
</tbody>
</table>

*Eigenvalues*  
4.15 2.72

*% Variance Explained*  
34.57 22.70

*Items were added to Learning from Errors subscale to reflect direct manipulation*
Table 6

*Manipulation Check Means and Standard Deviations by Condition*

| Measure                  | Error Management |   | Error Avoidant |   | Exploratory |   |  
|--------------------------|------------------|---|----------------|---|-------------|---|------
|                          | Mean | SD  | Mean | SD  | Mean | SD  | F        |
| Learning from Errors (T1)| 5.14 | 1.34 | 4.75 | 1.11 | 5.08 | 1.19 | 1.81     |
| Learning from Errors (T2)| 5.50 | 1.11 | 5.22 | 1.25 | 5.09 | 1.25 | 1.87     |
| Error Strain (T1)         | 3.45 | 1.46 | 4.23 | 1.56 | 3.81 | 1.46 | 4.46*    |
| Error Strain (T2)         | 2.81 | 1.35 | 3.23 | 1.58 | 2.76 | 1.29 | 2.01     |

N = 181; * p < .01

Despite partial support for the manipulation check measures, I still was confident that the manipulation was having its intended psychological effect for two reasons. One, manipulation effects using similar laboratory tasks have been quite small in past studies (Keith & Frese, 2005; Mangos & Steele-Johnson, 2001). Plus, I could argue that, at the very least, responses to the Learning from Errors measure exhibited marginal significance (p < .10) and were in the expected directions. Two, it is not surprising that the manipulation weakened with task experience. I would expect participants to rely on prior task experience as time increased and less on prior interventions.

**Tests of Hypotheses**

To test H1, I conducted a one-way ANOVA, with training type as the categorical, independent variable and average number of schedules and number of errors in the transfer trials...
as the dependent variable. In $H_1$, I predicted that training type would affect transfer performance. Specifically, I expected EMT participants to perform better than EAT participants on the transfer trials. I assessed performance in two different ways. One, I used number of completed schedules averaged across the two trials as an indicator of performance. Two, I averaged the number of errors committed across the two trials as a second indicator of performance. The former performance indicator, number of schedules produced, may be confounded with errors because participants can submit a schedule with errors. Thus, performance based on schedules produced may not be the best indicator of performance.

In the first test I used average number of schedules as the outcome and found that training type did not predict number of schedules produced in the transfer trial, $F(2, 178) = 0.14, p = .87$ (for means and standard deviations, see Table 7).

The second test used average number of errors committed on the transfer trials. Again, training type did not predict performance, $F(2, 178) = 0.29, p = .75$.

Even though the omnibus ANOVA tests were non-significant, I was more interested in the comparison between EMT and EAT groups for Hypotheses 1 through 4. Therefore, I directly compared these two groups with a one-way ANOVA. Expectedly, EMT participants did not produce more schedules on the transfer trials than EAT participants, $F(1, 127) = 0.229, p = .63$, nor did they produce significantly less errors than EAT participants, $F(1, 127) = 0.487, p = .49$. Thus, $H_1$ was not supported.

To assess $H_2$, $H_3$, and $H_4$, I used the Baron and Kenny (1986) approach to testing mediation. To establish partial mediation, I would need to show that the independent variable (i.e., training type) predicts the dependent variable (i.e., performance). Second, training type would need to predict the mediator. Third, the mediator must predict performance. Finally, the
training type-performance relationship must be reduced when controlling for the mediator. If the relationship is still significant, there is support for partial mediation. If the relationship is reduced to zero when with the mediator, then there are support for deducing complete mediation. In addition, I used the responses from Time 1 (after Trial 4) to predict performance within the transfer trials (Trials 5 and 6). For a summary of these analyses, see Tables 8 and 9.

In H2 I predicted that EMT participants would exhibit more meta-cognitive activities than EAT participants, which in turn would lead to greater performance effects. To assess this relationship, I used a one-way ANOVA, with training type as the categorical, independent variable, and average number of schedules and number of errors in the transfer trials as the dependent variable. Unfortunately, as seen in the results for H1, training type did not predict performance for either average number of schedules, \( F(2, 178) = 0.14, p = .87 \), or for average number of errors during the transfer trials, \( F(2, 178) = 0.29, p = .75 \). Therefore, we can not establish any form of mediation because a major assumption in the Baron and Kenny (1986) approach to testing mediation implies that the independent variable must predict the dependent variable. However, I will report the results for the other relationships for possible indirect effects.

Training type did not predict meta-cognition, \( F(2, 178) = 0.86, p = .43 \), and meta-cognition itself did not predict average number of schedules, \( \beta = -.05, t(1) = -0.67, p = .50 \), or average number of errors, \( \beta = .10, t(1) = 1.34, p = .18 \). A direct comparison between EMT and EAT participants showed that their levels of meta-cognition did not differ significantly, \( F(1,127) = 1.11, p = .29 \). Thus, H2 was not supported.

H3 stated that EMT participants would exhibit more emotion control than EAT participants, which in turn would lead to greater performance effects. Training type did not
predict emotion control, $F(2, 178) = 0.48, p = .62$, and emotion control did not predict average number of schedules, $\beta = .05, t(1) = 0.68, p = .50$, or average number of errors, $\beta = -.13, t(1) = -1.78, p = .07$. A direct comparison between EMT participants and EAT participants showed that they did not differ significantly in emotion control, $F(1, 127) = 0.99, p = .32$. Thus, $H_3$ was not supported.

Table 7

*Means and Standard Deviations by Condition on Various Outcomes*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Error Management</th>
<th>Error Avoidant</th>
<th>Exploratory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Average Transfer Schedules</td>
<td>4.49</td>
<td>2.10</td>
<td>4.32</td>
</tr>
<tr>
<td>Average Transfer Errors</td>
<td>1.68</td>
<td>1.60</td>
<td>1.46</td>
</tr>
<tr>
<td>Meta-Cognition (T1)</td>
<td>4.95</td>
<td>1.30</td>
<td>4.71</td>
</tr>
<tr>
<td>Emotion Control (T1)</td>
<td>4.30</td>
<td>0.64</td>
<td>4.18</td>
</tr>
<tr>
<td>Challenge Appraisals (T1)</td>
<td>4.83</td>
<td>1.23</td>
<td>4.70</td>
</tr>
<tr>
<td>Threat Appraisals (T1)</td>
<td>2.57</td>
<td>1.05</td>
<td>2.96</td>
</tr>
</tbody>
</table>

* $p < .10$

$H_{4a}$ stated that EMT participants would appraise the task as more challenging than EAT participants. This, in turn, would lead to greater performance effects. Training type did not predict challenge appraisals, $F(2, 178) = 1.67, p = .19$, and challenge appraisal itself did not predict average number of schedules, $\beta = -.09, t(1) = -1.27, p = .21$, or average number of errors,
\( \beta = -.01, t(1) = -0.18, p = .86. \) A direct comparison between EMT and EAT participants revealed that they did not differ significantly in appraising the task as challenging, \( F(1, 127) = 0.358, p = .55. \) Thus, \( H_{4a} \) was not supported.

\( H_{4b} \) stated that EAT participants would appraise the task as more threatening than EMT participants, which would in turn lead to greater performance effects on the transfer trials.

Training type marginally predicted threat appraisals, \( F(2, 178) = 2.75, p = .066. \) This

Table 8
*Summary of Main Effect and Mediation Analyses for Meta-Cognition, Emotion Control, Threat and Challenge Appraisals with Transfer Schedules as Outcome*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Predictor</th>
<th>( F/t )</th>
<th>( p )</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Type – Performance Relationships</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>Training Type</td>
<td>0.14</td>
<td>.87</td>
<td>N/A</td>
</tr>
<tr>
<td>Training Type – Mediator Relationships</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meta-Cognition</td>
<td>Training Type</td>
<td>0.86</td>
<td>.43</td>
<td>N/A</td>
</tr>
<tr>
<td>Emotion Control</td>
<td>Training Type</td>
<td>0.48</td>
<td>.62</td>
<td>N/A</td>
</tr>
<tr>
<td>Challenge Appraisals</td>
<td>Training Type</td>
<td>1.67</td>
<td>.19</td>
<td>N/A</td>
</tr>
<tr>
<td>Threat Appraisals</td>
<td>Training Type</td>
<td>2.75^</td>
<td>.06</td>
<td>N/A</td>
</tr>
<tr>
<td>Mediator – Performance Relationships</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>Meta-Cognition</td>
<td>-0.67</td>
<td>.50</td>
<td>- .05</td>
</tr>
<tr>
<td>Performance</td>
<td>Emotion Control</td>
<td>0.68</td>
<td>.50</td>
<td>.05</td>
</tr>
<tr>
<td>Performance</td>
<td>Challenge Appraisals</td>
<td>-1.27</td>
<td>.21</td>
<td>- .09</td>
</tr>
<tr>
<td>Performance</td>
<td>Threat Appraisals</td>
<td>-3.03*</td>
<td>.01</td>
<td>- .22</td>
</tr>
</tbody>
</table>

\(^p < .10, * p < .01\)
relationship was in the expected direction, as EAT participants appraised the task as more threatening than either EMT or EXP participants (see Table 3 for means and standard deviations). Threat appraisals also significantly predicted average number of schedules, $\beta = -0.22$, $t(1) = -3.03, p < .01$, but not average number of errors, $\beta = 0.00$, $t(1) = 0.01, p = .99$. Despite some positive support for some of the relationships within this mediation model, it is still unreasonable to test for mediation, given that training type did not predict performance.

Table 9
*Summary of Main Effect and Mediation Analyses for Meta-Cognition, Emotion Control, Threat and Challenge Appraisals with Transfer Errors as Outcome*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Predictor</th>
<th>$F/t$</th>
<th>$p$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Type – Performance Relationships</td>
<td>Training Type</td>
<td>0.29</td>
<td>.75</td>
<td>N/A</td>
</tr>
<tr>
<td>Training Type – Mediator Relationships</td>
<td>Training Type</td>
<td>0.86</td>
<td>.43</td>
<td>N/A</td>
</tr>
<tr>
<td>Meta-Cognition</td>
<td>Training Type</td>
<td>0.48</td>
<td>.62</td>
<td>N/A</td>
</tr>
<tr>
<td>Emotion Control</td>
<td>Training Type</td>
<td>1.67</td>
<td>.19</td>
<td>N/A</td>
</tr>
<tr>
<td>Challenge Appraisals</td>
<td>Training Type</td>
<td>2.75$^*$</td>
<td>.06</td>
<td>N/A</td>
</tr>
<tr>
<td>Threat Appraisals</td>
<td>Training Type</td>
<td>0.01</td>
<td>.99</td>
<td>.00</td>
</tr>
</tbody>
</table>

$^*$ $p < .10$
Nevertheless, there may be some claim for effects of training type on threat appraisals, so there is partial support for H$_{4b}$.

To test H$_{5}$ and H$_{6}$, I used a general linear model approach to test for moderation. Specifically, as predictors, I entered the training type variable and mean-centered personality variable in the first step and added the interaction term at the second step to predict average number of schedules and errors in separate analyses. I mean-centered the personality variables to reduce multicollinearity between model factors (Aiken & West, 1991). For summaries of the moderation analyses, see Tables 10, 11 and 12.

For H$_{5a}$ and H$_{5b}$, I expected adaptive perfectionism to be positively associated with transfer performance and maladaptive perfectionism to be negatively associated with transfer performance. Adaptive perfectionism was not related to average number of schedules, $F(1, 179) = 0.03, p = .86$, or average number of errors $F(1, 179) = 0.37, p = .54$, nor was maladaptive perfectionism, $F(1, 179) = 0.00, p = .97, F(1, 179) = 0.02, p = .89$. Thus, H$_{5a}$ and H$_{5b}$ were not supported.

For H$_{5c}$, I would expect adaptive perfectionism to interact with training type effects on transfer performance such that adaptive perfectionism would be positively related to the EXP and EAT condition but not related to EMT. Entering training type and mean-centered adaptive perfectionism in the first step, this analysis was not significant for either average number of schedules, $F(3, 177) = 0.10, p = .96$, or number of errors in the transfer trials, $F(3, 177) = 0.30, p = .82$. Adding the interaction term did not help the model achieve statistical significance for number of schedules, $F(4, 176) = 0.72, p = .58$, or errors, $F(4, 176) = 0.23, p = .92$. Thus, H$_{5c}$ was not supported.
Table 10
*Summary of Moderation Analyses for Adaptive Perfectionism*

<table>
<thead>
<tr>
<th>Model and Variable</th>
<th>$F$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transfer Schedules</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1:</td>
<td>0.10</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>Training Type</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2:</td>
<td>0.72</td>
<td>.016</td>
<td>.014</td>
</tr>
<tr>
<td>Training Type</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT x Personality</td>
<td>2.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Transfer Errors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1:</td>
<td>0.30</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td>Training Type</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2:</td>
<td>0.23</td>
<td>.005</td>
<td>.000</td>
</tr>
<tr>
<td>Training Type</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT x Personality</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 11
*Summary of Moderation Analyses for Maladaptive Perfectionism*

<table>
<thead>
<tr>
<th>Model and Variable</th>
<th>$F$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transfer Schedules</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1:</td>
<td>0.09</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>Training Type</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2:</td>
<td>0.08</td>
<td>.002</td>
<td>.000</td>
</tr>
<tr>
<td>Training Type</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT x Personality</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Transfer Errors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1:</td>
<td>0.21</td>
<td>.003</td>
<td></td>
</tr>
<tr>
<td>Training Type</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2:</td>
<td>0.34</td>
<td>.008</td>
<td>.005</td>
</tr>
<tr>
<td>Training Type</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT x Personality</td>
<td>0.74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For H5d, I would expect maladaptive perfectionism to interact with training type effects on transfer performance such that maladaptive perfectionism would be negatively related in the EXP and EAT condition, but not related in the EMT condition. Entering training type and mean-centered maladaptive perfectionism in the first step, this analysis was not significant for either average number of schedules, $F(3, 177) = 0.09, p = .96$, or number of errors in the transfer trials, $F(3, 177) = 0.21, p = .89$. Adding the interaction term did not help the model achieve statistical significance for number of schedules, $F(4, 176) = 0.08, p = .99$, or errors, $F(4, 176) = 0.24, p = .92$. Thus, H5d was not supported.

To assess H6a, I expected optimism to be positively related to the average number of schedules and negatively related to number of errors and discovered that optimism was not related to schedules $F(1, 179) = 0.70, p = .41$ or errors $F(1, 179) = 2.41, p = .12$ in the transfer trials.

For H6b, I would expect optimism to interact with training type effects on performance such that optimism would be positively related to the EXP and EAT condition, but not related to EMT. Entering training type and mean-centered optimism in the first step, this analysis was not significant for either average number of schedules, $F(3, 177) = 0.30, p = .82$, or number of errors in the transfer trials, $F(3, 177) = 0.93, p = .43$. Adding the interaction term did not help the model achieve statistical significance for number of schedules, $F(4, 176) = 0.52, p = .72$, or errors, $F(4, 176) = 0.72, p = .58$. Thus, H6b was not supported.
Table 12
*Summary of Moderation Analyses for Optimism*

<table>
<thead>
<tr>
<th>Model and Variable</th>
<th>$F$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transfer Schedules</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1:</td>
<td>0.30</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td>Training Type</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2:</td>
<td>0.52</td>
<td>.011</td>
<td>.006</td>
</tr>
<tr>
<td>Training Type</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT x Personality</td>
<td>1.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Transfer Errors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1:</td>
<td>0.93</td>
<td>.016</td>
<td></td>
</tr>
<tr>
<td>Training Type</td>
<td>0.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>2.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2:</td>
<td>0.72</td>
<td>.016</td>
<td>.000</td>
</tr>
<tr>
<td>Training Type</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean-Centered Personality</td>
<td>2.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT x Personality</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion

Due to the dynamic nature of industry, employees are required to adapt at a much faster pace than in past decades. To account for this change, one organizational area that can improve the adaptability of employees is the realm of training. Training paradigms normally encompass an error-avoidant perspective aimed at maximizing performance in a training context. For simple tasks, this approach has been shown to be effective (Keith & Frese, 2008). For more complex, dynamic tasks, this approach may have shortcomings. For instance, when faced with structurally distinct versions of a task domain, one’s knowledge base may be insufficient to complete the new task at a satisfactory level. Alternative approaches, such as error-management training (EMT), may prove to be effective at preparing employees for encountering novel situations in transfer settings.

For this reason, I sought to investigate three issues related to EMT research. One, I sought to demonstrate the ability of EMT participants to outperform EAT participants on a transfer task, which was a replication of previous research. Two, I wished to re-investigate the mediating mechanisms of EMT effects on performance and introduced the possibility of an additional mechanism. Specifically, Keith and Frese (2005) demonstrated that meta-cognition and emotion control mediated training type effects on performance and I expected to replicate this finding. Also, I adapted the use of cognitive appraisals (Lazarus and Folkman, 1984), specifically threat and challenge appraisal, partially to explain EMT effects on performance. Finally, I wanted to explore the inconsistent results of past research in regards to personality-treatment interactions (Gully et al., 2002; Heimbeck et al., 2003). To do this, I adopted Mischel’s (1968; 1977) use of situation strength for explaining the differential effects of training type interactions with personality variables on performance. To account for the situational
strength of EMT, I added an additional condition that was termed exploratory learning (EXP). The EXP condition was designed to be a weak situation to allow more variance between individuals than would be noticed in a stronger situation (e.g., EMT). I also incorporated two personality variables that I deemed to be relevant for a training context. Thus, I expected perfectionism and optimism to interact with EXP and EAT conditions but not with the EMT condition.

I will begin the discussion with three contributions of my study to the literature. Second, I will address five issues highlighted by this study that are likely to stir debate in the literature. I then will discuss the implications for future research based on these issues. On the basis of the theoretical issues presented, I will offer advice on the practical implications of this study, address this study’s potential limitations, and present my conclusion.

Contributions

I obtained partial support for $H_{4b}$, which stated that EAT participants would exhibit more threatening appraisals than EMT participants, and these appraisals in turn would have a negative effect on performance. On the basis of experimenter observation and open-ended debriefing, EAT participants found it to be more stressful to make errors because of the deliberate instructions to avoid making errors. If I had used a more liberal test of mediation (e.g., Sobel, 1988), as opposed to the conservative Baron and Kenny (1986) approach, I may have found full or partial mediation. Because I was unable to find a direct effect of training type on performance, I could not conclude that threat appraisal mediated the relationship.

Nonetheless, what does this mean for the constructs of emotion control and meta-cognition, two mediators that have been tested previously (Keith & Frese, 2005)? I believe that there is a possibility to integrate cognitive appraisals with the meta-cognition and emotion
control constructs to further explain EMT effects. That is, threat, challenge, meta-cognition and emotion control are related. Meta-cognition and threat appraisal are the only two constructs that do not correlate with each other. Nonetheless, emotion control correlates with meta-cognition, challenge appraisal, and threat appraisal. Meta-cognition and emotion control correlate significantly as well. Surprisingly, threat and challenge appraisals correlate positively with each other. It is possible that participants were interpreting the term “challenge” as a negative, stressful event as opposed to an opportunity to gain competence from a difficult task.

Needless to say, cognitive appraisal theory offers a more comprehensive theoretical model for incorporating the constructs of meta-cognition and emotion control. I would expect challenge appraisals to be positively related to emotion control and meta-cognition and threat to be negatively related with emotion control and meta-cognition. Although meta-cognition and threat were unrelated in this study, I would expect threatening appraisals to lead to withdrawal and distraction from the task, which would inherently lead to less meta-cognitive processing. Because of these threat appraisals, individuals may have less affective control of their emotions, which would lead to performance decrements. Challenge appraisals are commonly associated with positive emotions and task persistence in face of difficulty (Skinner & Brewer, 2002). This could lead to more meta-cognitive processing and emotion control. It is possible to integrate meta-cognition and emotion control into stress appraisal theory (Lazarus & Folkman, 1984) to encompass a more comprehensive theoretical explanation for future EMT effects.

A second contribution of this study is the modest support I received for training type effects on the manipulation check measure. EAT participants seemed to experience more strain from errors than EMT participants and EMT participants marginally held a more positive view that one can learn from errors than EAT participants. This finding suggests that one can
manipulate error valences among trainees, even though the effects of training type on performance was non-significant. Future research may help clarify how to improve the performance component of participants.

On a related note, a third contribution is the independent effects of the error strain measure on performance. Thus, individuals, regardless of condition, who experience more strain from making errors tend to perform worse than individuals who are not bothered by the presence of errors. This finding supports the notion that taking a negative view towards errors (or an error-avoidant attitude) can, in fact, hinder performance.

Theoretical Issues and Future Research

Based on the negligible results of training type effects on transfer performance, there are five issues that I will address that could contribute to the literature. These issues that are raised can help explain shortcomings of the present study.

One issue relates to the idea of task simplicity (or task difficulty). It could be that the task was inappropriate for this particular training intervention. Keith and Frese (2008) stated that participants would not benefit from either EMT or EAT when the tasks are analogical from one setting to the next. In other words, tasks that are too simple and lack distinction will be well served by either type of training. If this is a case of task simplicity, then we would expect a ceiling effect for performance and little differences when accounting for cognitive ability. However, this does not seem to be the case. Using self-reported ACT scores as a proxy for cognitive ability, I found that ACT scores predicted average performance on the number of schedules produced across the two transfer trials. Thus, if the task were indeed simple, participants across the broad spectrum of general mental ability would perform similarly well, which does not seem to be the case.
Related to the idea of task complexity, it is possible that the distinction between the training and transfer task was too subtle. The only difference between the two tasks was the addition of one rule to the transfer trials. This addition showed a marked decrement in performance from Trial 4 to Trial 5. Thus, it proved to be more complex and difficult to perform. Increasing task complexity is one way of designing transfer tasks. Past EMT research, however, not only increased the complexity of the task but also altered the structural distinctiveness of the task within the training and transfer trials. For example, Dormann and Frese (1994) used an SPSS package as the program of choice. Participants engaged in various tasks ranging in complexity from simple to complex. For instance, during the training phase of the program, participants worked on three tasks such as forming a data-entry layout, entering data, and calculating a correlation. For the transfer phase, Dormann and Frese (1994) made participants repeat the third task in the training phase (calculating a correlation) but increased the difficulty for the final two tasks in the transfer phase. In addition to computing a correlation, participants computed a $t$ test and finally computed a two-factor model using regression. Engaging in different activities while using the same program was something that the class scheduling task may have been unable to accomplish given the limited nature of the program. The ability to vary features of the class scheduling task was limited and thus we may have been unable to detect the training type effect.

Future research should investigate the components of transfer tasks and what makes one type of transfer task more sensitive to the effects of one type of training versus another. Is it simply altering the complexity of the same task or altering the different types of activities within a task domain that can best capture the effects of different training types? Future research should examine this question.
A second issue relates to the idea of time. EMT should take more time than EAT, given the nature of exploration and finding the correct methods for completing the task on one’s own time. Participants, in this study, were told that they had ten minutes per trial to work on the task and the participants could monitor the time with the clock that was provided on the computer screen. The clock counted upwards and the program stopped automatically when ten minutes had passed. I had decided to give participants access to the clock so that they could monitor progress and allocate resources to the task given the time constraints. Time constraints are common in a work setting, but they also can be stressful. EMT participants need to be encouraged to explore the task, and to do so, they may need more time to do so. When under time pressure, they may be forced to abandon EMT principles in favor of short-term production at the expense of long-term learning. Perhaps I erred when allowing access to a time clock during the experiment. Instead of the intended purpose of proper resource allocation, participants may have used the clock as a warning that time was running up, which resulted in either inefficient performance or withdrawal from the task in late stages of a trial. Participants were encouraged to “do their best,” but this may have contradicted the induction of positive error valence and exploration. Time pressure also may have weakened the manipulation at later times when administering the manipulation check measure because individuals in an EMT condition may have downplayed the usefulness of errors and exploration.

Previous EMT studies had partially controlled for this problem. For instance, Frese et al. (1991) and Dormann and Frese (1994) had participants from both conditions perform the same task under the same time constraints (i.e., four-hour training trial). EAT participants finished the tasks quicker than EMT participants but were encouraged to practice the tasks until the time
limits had elapsed. Thus, participants had ample time to complete and explore the task within the Frese et al. (1991) and Dormann and Frese (1994) samples.

Because time may have been an added stressor on participants and an unnecessary restraint to the efficacy of the manipulation, future research should consider eliminating the clock format or extending the sessions to enable participants to get the full benefits of the training interventions.

A third issue relates to previous methodological operationalizations of EMT. I am reluctant to take Keith and Frese’s (2008) line of reasoning too literally. It seems apparent that the two tenets of EMT (i.e., exploration and positive framing heuristics) as proposed in their meta-analysis are not the only components necessary for successful training interventions for complex tasks. They may be helpful components in later phases of learning, but approaches such as enactive exploration (Wood et al., 2000) and delayed emotion control training (Bell & Forman, 2009) may be better suited to tackle some of the deficiencies that I had witnessed in this study. This would allow for more guided mastery training in early phases of training. This may enhance self-efficacy early in training, help build a basic knowledge base for future reference, and set up trainees for successful exploration at more complex stages of learning.

A sub-issue relates to the use of instruction manuals during training sessions. In an effort to reduce a potential confound in the experimental design prior to data collection, I made the decision to allow participants the opportunity to review task instructions when playing the audio recording of the instructions. However, they did not have access to the instructions when working on the task. By preventing participants from having access to instruction manuals during training, all participants had the same amount of time to look at the instructions. This also may have contributed to the lack of training type effect. Most, if not all, previous EMT
studies allowed participants to have access to instruction manuals during training phases but not during test phases (Chillarege et al., 2003; Frese, et al., 1991; Dormann & Frese, 1994; Heimbeck et al., 2003; Keith & Frese, 2005). The question that arises then, is whether participants relying on exploration or manual consultation? There is a lack of statistical or experimental control for the instructions in previous studies, even though each training type has equal access to the instructions. This may make it difficult to interpret previous findings.

There is one possibility to account for the use of instruction manuals during training and to overcome the shortcomings of previous studies. One can run a 2 x 2 between-subjects factor study where one factor is training type and the other instruction manual access. Thus, there would be four cells and they would be mutually exclusive to participants. One factor would be training type (with EMT and EAT as the two conditions) and instruction manuals as the other between-subjects factor (with or without instruction manuals during training phases of task practice). I would expect participants with access to training manuals to perform better than those without access to said manuals in the beginning of training, which would not be very surprising. As a reminder, participants would all receive the same introductory audio-recorded instructions prior to the first trial of practice. I also would expect EMT participants with access to manuals would perform better on later trials given their ability to gain baseline knowledge early in training in addition to the propensity to explore on later trials. They seemingly would perform better than EMT and EAT participants without access to instructions (who presumably would not differ on performance based on the results from this study) because of the instructions factor. EMT participants with instructions also would perform better than EAT participants with instructions on later trials but not early trials, which would corroborate prior research. In case experimental control did not work as intended, it might be advisable to statistically control for
exploratory behavior as well as manual consultation through experimenter observation. This type of design could overcome some of the problems of prior research and the current study.

Another possibility to improve the current study is to use forced-error sessions. As described earlier, Nordstrom et al. (1998) and Chillarege et al. (2003) gave participants rudimentary training and then put them through an error situation and differentiated the conditions in how they could deal with the errors. This is akin to Wood et al.’s (2000) conceptualization of enactive exploration. Perhaps this could be explored in future research.

Dormann and Frese (1994) offered support for the use of instruction manuals during training phases. Instead of counting the amount of times participants consulted their instruction manuals, they measured the amount of exploration in both conditions with the assumption that trainees will explore a task domain despite trainer recommendations to avoid errors. Dormann and Frese (1994) investigated the effects of training type on performance using a word-processing program. EMT participants outperformed EAT participants on the transfer tests, as hypothesized. The researchers also separated the conditions according to high and low exploration as rated by external observers. The authors hypothesized that high exploration would be more beneficial than low exploration for the performance test regardless of condition. Indeed, even in the EAT condition, they found that participants who explored more benefited from the exploration. The same was found in the EMT condition. Despite the inability to account for the potential confound of the instruction manual, Dormann and Frese (1994) argued that participants who explored regardless of condition benefited from its use. Assuming those low in exploratory behavior were the ones who consulted the instruction manuals more often than not, the above finding strengthened their argument that exploration was key to gaining knowledge and performing better on transfer tasks. Perhaps the use of instruction manuals early
on in training provided a good conceptual basis for participants to understand the fundamental aspects of the task domain, and later exploration aided in understanding more complex aspects of the domain. Bell and Forman (2009) would support this assertion.

Wood et al. (2000) would suggest that enactive exploration would be most beneficial for trainees because they are given guided mastery during the initial phases of training and then are encouraged to explore the task. Other researchers, such as Nordstrom et al. (1998) and Chillarege et al. (2003), had introduced a forced error session where participants in both conditions experience errors for the first time. Previous sessions are error-free so that they can learn task fundamentals. Once the forced error session occurs, EAT participants receive step-by-step instructions to solve the problem whereas EMT participants are told to explore to find the answer. Participants then are tested in performance phases. Future research should continue to explore this avenue.

Future research could also explore the use of alternative outcomes or measures of outcomes. For instance, what types of errors affect learning? Are people making one error multiple times or five different errors? One also could measure schemas directly or use the distinction of learners and non-learners for assessing the effectiveness of training interventions.

A fourth issue relates to the comparison of training type conditions and the respective psychological perceptions (i.e., meta-cognition, emotion control, challenge). Most of the relationships between various person-task perceptions and training type were small and non-significant. A possible explanation for this could rest with the combination of reasons for the failure of $H_1$ and past research on person-task perceptions. Based on the results of the three previous mediation tests, it is evident that participants are behaving slightly in response to the manipulation, given the directionality of responses to the mediation measures and given the
positive responses to the manipulation check measure (first administration). Despite this evidence, we do not see significant effects because other factors may influence performance, such as cognitive ability, task experience, etc. Past research on self-efficacy has shown that future efficacy assessments are influenced by past performance (e.g., Heggested & Kanfer, 2005; Shea & Howell, 2000). If all participants across conditions perform similarly, as was the case in this study, then they all would achieve similar task perceptions (i.e., emotion control, challenge, etc.). Future research then should attempt to experimentally manipulate EMT person-task perceptions, such as meta-cognition, cognitive appraisals, and emotion control. This would allow researchers to disentangle the effects of past performance on future perceptions of performance. Also, I am optimistic that future studies that rectify the possible design problems of EMT research could obtain greater response discrepancies on the self-report measures across conditions.

A final issue concerns the effects of personality on performance. It could be that certain personality variables are only appropriate or relevant to particular task domains. For instance, Barrick, Mount, and Judge (2001) found that conscientiousness is the best predictor of job performance because according to the work domain, being achievement-oriented, organized and timely is beneficial for this situation. However, agreeableness was not related to sales performance at all. Perhaps it was the case that performing the class scheduling task had no relevance for optimists or perfectionists. A related issue is that participants relied on their cognitive capacities to perform the task. In the work domain, it is thought that personality exhibits its strongest effects on contextual factors, such as organizational citizenship behaviors, whereas cognitive factors affect task-based outcomes (Motowidlo, Borman, & Schmit, 1997). Under these assumptions, it is no surprise that I failed to find an effect of personality on
performance. Also, the performance measure was the average performance from Trial 5 and 6, so it is highly likely that by this point in the task, people were relying heavily on their previous experience for influencing their behavior instead of personality.

Given that attribute-treatment interactions have proved elusive in past EMT research (Gully et al., 2002; Heimbeck et al., 2003), it is possible that we should investigate more cognitively-based attributes, given Motowidlo et al.’s (1997) model. Future researchers should investigate this avenue.

Practical Implications

Given the implications of this study, it might be premature to advise organizations regarding the efficacy of EMT. Despite the positive results that Keith and Frese (2008) expound, the meta-analysis was based on a relatively small sample of studies ($k = 24$) on a limited number of tasks. What those results do tell us is that there are early signs of promise and that active learning with high personal control on complex computer-based tasks may benefit from an EMT intervention. However, the current study would suggest that there is still more work to be done. Also, training can be stressful for some participants, as seen in this study as some participants responded with greater threat appraisals and error strain. Neutralizing stressful sources can provide trainees with the means to accomplish successful learning during training, as was seen with the finding that EAT participants made more threat appraisals than EMT participants. As expected, threat hindered performance. Based on this evidence, error-avoidant attitudes during instruction may impede trainees from fully accomplishing what they could learn. Nonetheless, once the laboratory deficiencies are fully investigated, one can advise organizations on the practical implications that EMT may have in the workplace.

Limitations
One major limitation involved the experimental procedure. In attempting to strengthen
the manipulation, there may have been a disruption in task familiarization. To reiterate,
participants listened to audio-recorded task instructions while following along with the text
version of the recording. This process lasted approximately 7 minutes. Participants handed in
the instruction sheets and read the task rules. This took an additional 1-2 minutes of time.
Following the task rules, participants were read the manipulation induction and then instructed to
copy down the instructions while the experimenter resuscitated the induction line-by-line. This
took approximately 10 minutes to complete. Also, there was a feedback cuing screen that
participants read that differed across conditions, which took approximately 30 seconds.

The aforementioned activities that occurred between task instructions and task
performance may have led to retroactive interference. This occurs when learning activities at a
later time interferes with earlier-learned material when attempting to retrieve previous material.
This assumes that the material is available at times of retrieval, but not accessible because later-
learned material is interfering with the accessibility of the earlier-learned material. Thus, I can
not be certain that retroactive interference actually took place, given that I have no knowledge of
participants’ knowledge at that time.

A likelier scenario would be that there was no interference but that participants across
conditions engaged in homogeneous exploration to understand the task domain. Homogeneous
exploration simply means that individuals, regardless of condition, were forced to explore the
task features due to a possible lack of practice or lack of learning the initial instructions. One
way to test this assumption, admittedly not the best, is to analyze the number of times
participants accessed rule call ups at Time 1. (Note: Participants could request to review any
rule during each trial using the F# function.) Here, I found that there were no differences
between conditions in number of times that participants used the rule call-up functions, \( F(2, 178) = 0.08, p > .05 \). This test offered an objective means of analyzing the amount of exploration that occurred across conditions. This test suggests that all participants engaged in homogeneous exploration. There are three reasons to believe that homogeneous exploration occurred.

One, although the manipulation had the intended psychological effect in how they responded to the manipulation check measure, this effect did not translate into how they performed. Thus, participants may have been more concerned with understanding the task domain regardless of the manipulation.

Two, after administering the manipulation check measure the second time (after Trial 4), the responses to the manipulation check were much weaker and more homogeneous across conditions, suggesting that the effect of the manipulation had declined. This is not surprising given that as participants gain task experience, they will rely more heavily on previous task experiences than an induction administered towards the beginning of the session.

A final reason for believing that participants may have engaged in homogeneous exploration was that they were not able to practice what they learned from instructions immediately after listening to the audio-recording and thus failed to gain baseline knowledge of the task interface. Therefore, participants had to resort to whatever tactics they could to understand the domain. This would attenuate any manipulation that was being attempted. Other research suggests that the EMT induction may be better off delayed to Trial 3 at least to give participants enough time to establish baseline knowledge for novel stimuli (Bell & Forman, 2009; Heimbeck et al., 2003). Once individuals establish enough competence with a certain task they may be able to explore features in more complex ways.
Another limitation of the current study was the use of a student sample. The sample demographics make generalizability difficult and the use of students may hinder motivation to perform in a laboratory setting, which may obstruct researchers from fully investigating the constructs of interest. Highhouse and Gillespie (2009) suggested that we do not generalize samples or settings, but rather we generalize theories.

A related issue with the use of student samples is that, given the nature of the task, perhaps there was a possibility that participants were already familiar with performing the task of scheduling classes in everyday life. If this were the case, then different types of training would not contribute much to the learning of the task. On the other hand, participants were unfamiliar with this particular computer interface. The programming software is in MS-DOS format (outdated by today’s standards) and only has been used in laboratory settings, so some level of learning is required to perform the task. Plus, as I already have stated, cognitive ability plays a role in learning this task. In addition, employees in most job settings have at least a rudimentary idea of most task domains when entering a training session, so to suggest that all participants did not need training in this experiment given their prior experience does not hold as the following data suggests.

Seventy-two percent of the sample was college freshmen, who have had presumably only three quarters worth of experience scheduling classes. Twenty percent were college sophomores. The cumulative percentage of the sample who were freshmen and sophomores was 92.3%. Also, nearly half (49.2%) of the sample said that they had no prior experience with scheduling tasks such as the one they performed in the experiment. In addition, the amount of experience with similar tasks correlated with class rank, $r(180) = .197, p = .008$, such that the less experience one had with similar tasks, the greater likelihood that one was a freshman or sophomore. Stated
differently, the more task experience one has, the greater the likelihood that one is a junior or a senior. Task experience also does not predict performance, $F(1, 179) = 0.189$, $p = .66$. Thus, even if participants had task experience, it did not seem to matter much. I conclude that task domain familiarity was not a major limitation with this particular study.

One limitation with the mediation analyses was that I used the Baron and Kenny (1986) approach, which is a conservative test considering that if one does not establish a relationship between the independent variable and dependent variable, then one can not conclude there is mediation. As $H_1$ showed, there was no direct effect of training type on performance, so thus, no mediation. To that end, the Baron and Kenny (1986) method suffers from low power due to the strict testing guidelines (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). Sobel (1982) offers a more liberal test of mediation.

Stone-Romero and Rosopa (2008) also would contend that any test of mediation should use randomized experimental designs to test every aspect of the mediation model. Their main argument stems from the overuse of causal modeling using non-experimental data and concluding that one can make causal inferences. However, this study did not use causal modeling and further, I manipulated two of the relationships within the mediation model. The first was the effect of training type on performance as well as training type on the mediating variables. It is possible that future research could manipulate any of the mediating variables (Bell & Forman, 2009; Keith & Frese, 2005).

The proposed shortcoming in regards to meta-cognition could be due to poor measurement of the meta-cognition construct. For this study, I used a self-report measure adapted from a generalized stress and anxiety scale (Cartwright-Hatton & Wells, 1997). Participants answered this measure in a retrospective manner such that the experimenter
administered the measure after Trial 4 and Trials 6. This method could have been subjected to response bias or memory distortion. The original study that tested meta-cognition used a verbal protocol analysis to measure the construct, in which experimenters coded appropriate thought processes while participants worked on the task (Keith & Frese, 2005). Also, Keith and Frese (2005) experimentally manipulated the variable such that one of the two EMT conditions received instructions to “think aloud” whereas the other did not. It is possible that there was not an effect to be found or that the measure was too poor to detect an effect. Because previous research had found an effect, we would expect to observe one as well (Keith & Frese, 2005). However, the manipulation did not work as intended, so it is possible that in this particular study the effect did not occur. A possibility for future research is to pilot test the use of a retrospective self-report measure with a verbal protocol analysis to validate the use of a self-report measure. A potential problem with retrospective self-report measures of meta-cognition is the possibility for attribution bias. Some may report that they thought about their plans and strategies because they performed well, not whether they engaged in meta-thoughts or not. Proper pilot testing may help researchers decide on the validity of these types of measures. Self-report measures would ease the data collection process instead of the time-consuming nature of verbal protocol analysis.

Another limitation of the current study is the task itself. Previous research using this task have used it as a means of assessing motivational tendencies of participants such as goal orientation, self-efficacy, subjective task complexity, etc. (Mangos & Steele-Johnson, 2001; Steele-Johnson, Beauregard, Hoover, & Schmidt, 2000). Performance was measured by how many schedules were submitted. Provision of effort was deemed to be more important than the necessary quality of the schedules that were produced. These were reasonable accommodations given the nature of the constructs that were being explored. However, this task may have been
unsuitable for the purposes of the current study, where quality was more important to assess learning. Plus, participants in past studies may have been more interested in the tasks that they were learning (such as PowerPoint, Word, SPSS, etc.). These tasks would prove more useful for participants outside of the laboratory setting than the class scheduling task. I would advise future researchers to select their respective task carefully and pick ones that may have relevance for participants in their own lives.

Conclusion

Despite the null results of the hypotheses, there are two areas in which my results should enrich this area of research. First, researchers should not be content with Keith and Frese’s (2008) assertion that exploration and framing statements are the key components to successful EMT interventions. Based on this study, there are more components to uncover, especially in relation to the use of instruction manuals during training phases. In addition, other researchers are suggesting a hybrid approach to error-encouragement training (Bell & Forman, 2009; Wood et al., 2000). Second, I have offered a new theoretical framework in cognitive appraisal theory that has shown early signs of offering a more comprehensive structure for investigating EMT. These two main contributions can help further the development of EMT as an alternative training approach.
References


Appendix A

Task Rules

Training Task

1. A student’s schedule may not contain 2 different sections of the same class.
2. A student’s schedule may not contain 2 course sections scheduled for the same time and day.
   In other words, 5 course sections that you assign to a student’s schedule must meet at different, non-overlapping time periods.
3. Each student’s schedule must be unique. You may not assign the identical 5 course sections to 2 or more students. In other words, each student’s schedule must differ by at least one course section from other student’s schedules.
4. You must assign a lab section for any course in which a lab is required. Thus, if “Lab required” is displayed in the course schedule window for a course, you must schedule a lab section as well as the lecture section for that course.
5. You can not skip a student’s class plan. You must complete each class plan fully before you move on to the next.
6. Some students prefer classes on certain days (i.e., MWF or TTH). You can not schedule a class on days not preferred by the student.

Transfer Task

1. A student’s schedule may not contain 2 different sections of the same class.
2. A student’s schedule may not contain 2 course sections scheduled for the same time and day.
   In other words, 5 course sections that you assign to a student’s schedule must meet at different, non-overlapping time periods.
3. Each student’s schedule must be unique. You may not assign the identical 5 course sections to 2 or more students. In other words, each student’s schedule must differ by at least one course section from other schedules you have created.
4. You must assign a lab section for any course in which a lab is required. Thus, if “Lab required” is displayed in the course schedule window for a course, you must schedule a lab section as well as the lecture section for that course.
5. You can not skip a student’s class plan. You must complete each class plan fully before you move on to the next.
6. A student’s schedule must not exceed 3 consecutive hours of lecture time without having a break. Thus, if a student is scheduled for two hours of lecture, then you must allow that student at least 30 minutes of break time before the next scheduled class.
7. Some students prefer classes on certain days (i.e., MWF or TTH). You can not schedule a class on days not preferred by the student.
Appendix B

**Manipulation Description**

**Error Management Training**

When performing the task, you may come across obstacles that may force you to make mistakes. Some research has shown that the best learning method is through trial-and-error, so it is okay to make mistakes. Errors are a natural part of learning and aid in developing your understanding of the task. When you encounter an error, do not feel stressed or dwell on it. Feel free to explore the task and discover why you made that error. Understanding how and why the error occurred will help you improve your performance on this task as you continue.

**Exploratory Learning**

When performing the task, you may come across obstacles that may force you to make mistakes. Although the experimenter will not be able to assist you during the task, you should feel free to explore the task in order to correct problems that arise while working on the task. We encourage you to discover as much as you can about the nature of this task.

**Error Avoidant Training**

When performing the task, you may come across obstacles that may force you to make mistakes. You want to avoid mistakes at all costs. Errors force you to work inefficiently and decrease your performance. Make sure that you follow the instructions given to you. The rules and guides will help learning how to perform the task. Remember, errors are counted, so you want to try to avoid mistakes as much as possible.
Appendix C

Pilot Study Results

I had a three-fold purpose for conducting pilot research prior to the beginning of full data collection. First, I wanted to check the reliability of the meta-cognition and manipulation check measures. Specifically, the meta-cognition measure was adapted from a dispositional measure to reflect features of the task. There was no reason to check the reliability of the two subscales of the Error Orientation Questionnaire (i.e., the manipulation check measure) other than to recheck reliability of the measure used in the experiment. A second purpose for the pilot study was to see if the manipulation had the intended psychological effect on participants. Thus, I compared the responses to the manipulation check items by condition. The final reason for the pilot study was to check the expected direction of effects on performance by condition.

Pilot Study #1

For the meta-cognition measure, reported Cronbach alphas were sufficient in Trials 1-5 ($\alpha = .85, .86, .93, .92, \text{ and } .88$, respectively). Similarly, for the two subscales of the manipulation check, the Cronbach alphas for both the learning from errors subscale and the error strain subscale were deemed sufficient ($\alpha = .76 \text{ and } .83$, respectively).

For the learning from errors subscale, higher responses would signify a more positive view of errors. Thus, I expected participants in the EMT condition to score highest on this subscale, followed by those in the EXP condition, with those in the EAT condition having the lowest scores. Although EMT participants did score higher than participants from the other two conditions ($M = 5.83, SD = .99$), their scores were not significantly greater than scores observed for EAT ($M = 5.64, SD = 1.31$) or EXP participants ($M = 4.92, SD = 1.35$).
Appendix C, continued

For the error strain subscale, higher responses would signify a more negative view of errors. Thus, I expected EAT participants to score highest on this subscale, followed by EXP participants, with EMT participants reporting the lowest scores. Unfortunately, EMT ($M = 4.68, SD = 1.03$) and EAT ($M = 4.69, SD = .84$) participants responded similarly to this subscale whereas EXP participants scored the lowest ($M = 3.57, SD = 1.04$).

Responses to open-ended debriefing questions following the task simulation may indicate reasons for the lack of differential responses to the manipulation check measure. Participants expressed the perception that the subscales were measuring dispositional personality tendencies. In other words, after the manipulation, participants read the manipulation check items as assessing perceptions of attitudes towards errors in everyday life. Thus, the measure did not capture any psychological induction that may have been present within the experiment. For this reason, there was a need for a second pilot study. I will discuss changes to the protocol in the next section titled, *Pilot Study #2*.

Before I describe the second pilot study, I will report performance and error counts for the first pilot study. Despite the fact that there was no evidence that the manipulation had the intended psychological effects on participants, I elected to report the outcomes because it represents the third purpose of the pilot research.

Table 1 displays the performance means by trial. The first four trials represent training trials whereas the fifth trial is the lone transfer trial. Based on the manipulation check results, I would expect to see an effect on number of schedules produced per trial to be only due to chance. As expected, there were no significant differences between conditions across trials, $F (2, 20) = 0.500, p = .614$. In Trial 2, EXP participants outperformed EMT and EAT participants by
Appendix C, continued

Table 1

*Average Number of Schedules per Trial by condition (Pilot Study #1)*

<table>
<thead>
<tr>
<th>Trial</th>
<th>Condition</th>
<th>EMT (n = 10)</th>
<th>EAT (n = 9)</th>
<th>EXP (n = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>3.50</td>
<td>2.09</td>
<td>2.78</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>8.90</td>
<td>1.79</td>
<td>8.22</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>12.20</td>
<td>2.70</td>
<td>10.44</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>14.00</td>
<td>3.92</td>
<td>13.67</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>4.10</td>
<td>2.08</td>
<td>4.33</td>
</tr>
</tbody>
</table>

approximately 1.5 to 2 schedules, respectively. However, this was not in the expected direction. We would expect EAT participants to perform best during the training trials followed by EMT and EXP participants. By the fifth trial, EMT participants should have the highest scores, followed by EXP and EAT participants. As seen in Table 1, there are either no differences, differences due to chance, or differences not in the expected direction. Plus, the sample size is quite small (e.g., n = 6 for EXP condition).

The same could be said in regard to the outcome of error counts. I would expect to observe more errors committed by EMT participants in early trials than by EXP or EAT participants. Also, by the fifth trial, I would expect EMT participants to account for the fewest
Appendix C, continued

errors, compared with other participants. Indeed, as shown in Table 2, EMT participants committed on average 1.78 errors compared with EXP and EAT participants ($M = .40$ and .44, respectively) in the first trial. However, the observed pattern of means led me to conclude that there were no discernible differences between conditions on later trials. Based on the manipulation check measure responses, we can not assume that participants are behaving in accordance with the manipulation, and thus any results could be due simply to chance.

Pilot Study #2

To address the manipulation check problem found in the first pilot study, I made two alterations to the measure itself. First, I altered the items to reflect task-specificity in order to cue participants into thinking about how they currently view errors as opposed to how they generally view errors in normal life circumstances. Second, based on the recommendation of an

<table>
<thead>
<tr>
<th>Trial</th>
<th>EMT ($n = 9$)</th>
<th>EAT ($n = 9$)</th>
<th>EXP ($n = 5$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
</tr>
<tr>
<td>1</td>
<td>1.78  3.03</td>
<td>0.44  0.53</td>
<td>0.40  0.55</td>
</tr>
<tr>
<td>2</td>
<td>0.33  0.50</td>
<td>0.11  0.33</td>
<td>0.20  0.45</td>
</tr>
<tr>
<td>3</td>
<td>0.11  0.33</td>
<td>0.78  2.33</td>
<td>0.20  0.45</td>
</tr>
<tr>
<td>4</td>
<td>0.33  0.71</td>
<td>0.56  1.13</td>
<td>0.20  0.45</td>
</tr>
<tr>
<td>5</td>
<td>1.33  1.50</td>
<td>1.33  1.23</td>
<td>1.60  1.34</td>
</tr>
</tbody>
</table>
Appendix C, continued

advisor, I added three items that were more direct in asking about the manipulation itself (see Appendix D for the appropriate items).

I made two changes to strengthen the manipulation. One, instead of one administration of the manipulation check measure directly after the induction, I administered the revised measure twice: once after Trial 1 and once after Trial 4. The second alteration coincided with a change in the study protocol to administer the manipulation text twice during the experiment, with the experimenter reading the text aloud. This procedure contrasted with Pilot Study #1 in which the manipulation was administered once and participants read the text of the manipulation on a computer screen and then copied down the content onto a sheet of paper. Open-ended debriefing after Pilot Study #1 suggested that having an experimenter read the paragraph aloud may enhance memory and attention to the induction as opposed to participants reading the induction on a computer screen. To ensure encoding of the information, the text was read aloud once in its entirety, followed by a line-by-line recitation of the same text. The experimenter gave enough time between recitations for participants to write down the paragraph. Finally, the experimenter reiterated the paragraph once more in its entirety after Trial 4 (participants did not copy down the paragraph at this time).

I also made two alterations to the task. One, based on participant recommendations in Pilot Study #1, we made the task more difficult by adding one rule to both the training and transfer trials. Participants expressed a lack of challenge in previous task versions. Instead of five rules for the training trials, there were now six rules. Instead of six rules for the transfer trials, there were now seven rules (specific rules can be found in Appendix A). Also, we added a second transfer trial for a total of six, 10-minute trials.
Appendix C, continued

The Cronbach alphas for the task-specific measures of learning from errors and error strain subscales at Time 1 are .93 and .66, respectively. At Time 2, the learning from errors subscale had a reliability of .89 whereas the error strain subscale reliability was .80.

Now that I had assessed the reliability of the revised measures, I proceeded to evaluate the revised manipulation check measures by condition. Means of the manipulation check responses can be seen in Table 3. Although responses to the revised manipulation check measures (i.e., learning from errors and error strain subscales) were in the expected direction, there were no significant differences between the conditions, $F(2, 53) = 1.22$, $F(2, 53) = 0.38$, $F(2, 53) = 0.46$, $F(2, 53) = 0.32$, all $p$’s $> .05$ (learning from errors at Time 1, learning from errors at Time 2, error strain at Time 1, error strain at Time 2, respectively). With respect to the additional three items, all were in the expected directions at both times of administrations. However, only one item (MC 2 at Time 2) was statistically significant, $F(2, 53) = 3.76$, $p < .05$. Nonetheless, given the small sample size and based on open-ended debriefing questions, I felt confident that the manipulation was having the desired effect. Therefore, I proceeded to analyze the performance outcomes.

Tables 4 and 5 display the average number of schedules produced and errors committed by trial. On most trials, EAT participants tended to outperform EMT and EXP participants. However, given that EAT participants tended to make more errors than EMT participants, the number of schedules may be confounded with more errors. A primary precept of error management training is to ensure that fewer errors are made on transfer tasks. Given that EMT participants committed fewer errors on average in the transfer phase of the experiment than EAT
Appendix C, continued

participants, $t (53) = 1.97, p < .05$, I concluded that the manipulation was having its intended psychological effect on performance.

Table 3

*Manipulation Check Means by Condition (Pilot Study #2)*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Condition</th>
<th>EMT ($n = 22$)</th>
<th>EAT ($n = 18$)</th>
<th>EXP ($n = 16$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Learning from Errors (Time 1)</td>
<td></td>
<td>5.63</td>
<td>1.30</td>
<td>5.40</td>
</tr>
<tr>
<td>Learning from Errors (Time 2)</td>
<td></td>
<td>5.64</td>
<td>1.22</td>
<td>5.54</td>
</tr>
<tr>
<td>Error Strain (Time 1)</td>
<td></td>
<td>3.56</td>
<td>0.98</td>
<td>3.90</td>
</tr>
<tr>
<td>Error Strain (Time 2)</td>
<td></td>
<td>2.94</td>
<td>1.24</td>
<td>3.28</td>
</tr>
<tr>
<td>MC 1 (Time 1)</td>
<td></td>
<td>5.32</td>
<td>1.73</td>
<td>5.00</td>
</tr>
<tr>
<td>MC 1 (Time 2)</td>
<td></td>
<td>5.95</td>
<td>1.73</td>
<td>5.39</td>
</tr>
<tr>
<td>MC 2 (Time 1)</td>
<td></td>
<td>4.05</td>
<td>1.59</td>
<td>3.06</td>
</tr>
<tr>
<td>MC 2 (Time 2)</td>
<td></td>
<td>4.59</td>
<td>1.84</td>
<td>3.22</td>
</tr>
<tr>
<td>MC 3 (Time 1)</td>
<td></td>
<td>5.05</td>
<td>1.79</td>
<td>4.33</td>
</tr>
<tr>
<td>MC 3 (Time 2)</td>
<td></td>
<td>5.45</td>
<td>1.60</td>
<td>4.39</td>
</tr>
</tbody>
</table>
### Table 4

**Average Number of Schedules per Trial by condition (Pilot Study #2)**

<table>
<thead>
<tr>
<th>Trial</th>
<th>EMT (n = 22)</th>
<th>EAT (n = 18)</th>
<th>EXP (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>1</td>
<td>1.91</td>
<td>1.85</td>
<td>1.78</td>
</tr>
<tr>
<td>2</td>
<td>3.86</td>
<td>2.66</td>
<td>4.89</td>
</tr>
<tr>
<td>3</td>
<td>4.68</td>
<td>2.33</td>
<td>6.11</td>
</tr>
<tr>
<td>4</td>
<td>6.09</td>
<td>2.86</td>
<td>6.56</td>
</tr>
<tr>
<td>5</td>
<td>4.14</td>
<td>1.91</td>
<td>3.94</td>
</tr>
<tr>
<td>6</td>
<td>4.45</td>
<td>2.35</td>
<td>5.39</td>
</tr>
<tr>
<td>Avg. 1-4</td>
<td>4.14</td>
<td>2.05</td>
<td>4.83</td>
</tr>
<tr>
<td>Avg. 5-6</td>
<td>4.30</td>
<td>1.76</td>
<td>4.67</td>
</tr>
</tbody>
</table>
Appendix C, continued

Table 5

*Average Number of Errors per Trial by condition (Pilot Study #2)*

<table>
<thead>
<tr>
<th>Trial</th>
<th>EMT (n = 22)</th>
<th>EAT (n = 18)</th>
<th>EXP (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
</tr>
<tr>
<td>1</td>
<td>3.27 4.51</td>
<td>3.61 4.64</td>
<td>2.12 2.92</td>
</tr>
<tr>
<td>2</td>
<td>0.50 1.14</td>
<td>0.61 1.65</td>
<td>0.81 2.32</td>
</tr>
<tr>
<td>3</td>
<td>0.55 0.96</td>
<td>1.28 2.42</td>
<td>0.75 1.61</td>
</tr>
<tr>
<td>4</td>
<td>0.55 1.10</td>
<td>0.61 1.04</td>
<td>1.19 2.32</td>
</tr>
<tr>
<td>5</td>
<td>1.14 1.49</td>
<td>1.67 1.78</td>
<td>0.88 0.89</td>
</tr>
<tr>
<td>6</td>
<td>0.68 0.78</td>
<td>1.50 1.62</td>
<td>1.62 1.71</td>
</tr>
<tr>
<td>Avg. 1-4</td>
<td>1.22 1.41</td>
<td>1.53 1.74</td>
<td>1.22 1.24</td>
</tr>
<tr>
<td>Avg. 5-6</td>
<td>0.91 0.92</td>
<td>1.58 1.32</td>
<td>1.25 0.97</td>
</tr>
</tbody>
</table>
Appendix D

Manipulation Check

To what extent do the following statements apply to you:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly agree</td>
</tr>
</tbody>
</table>

**Error Strain**
1. I find it stressful when I make an error.
2. I am often afraid of making mistakes.
3. I feel embarrassed when I make an error.
4. If I make a mistake, I “lose my cool” and become angry.
5. While working, I am concerned that I could do something wrong.

**Learning from Errors**
1. Mistakes assist me to improve my performance.
2. Mistakes provide useful information for me.
3. My mistakes help me to improve my performance.
4. My mistakes have helped me to improve my performance.
5. While working on the task, I was able to maintain task focus.*
6. I felt positive after making a mistake on the task.*
7. While working on the task, I was encouraged to explore the task.*

*Scoring Procedures:* Item responses are summed to create aggregate scores on the particular subscale. Divide the Error Strain score by 5 and the Learning from Errors score by 4 to obtain score in its original metric. Higher responses on the Learning from Errors subscale signifies a positive view towards errors whereas higher score on the Error Strain subscale would signify an aversion towards making errors.

* Added items
Appendix E

Demographics

Please provide the best possible answer that corresponds to you.

1. What is your age?
   __________ years of age

2. What is your gender?
   1 Male  2 Female

3. What is your class rank?
   1 Freshman  2 Sophomore  3 Junior  4 Senior  5 Other

4. What is your major?
   1 Business  2 Communications  3 Education  4 Engineering
   5 Mathematics  6 Psychology  7 Sociology  8 Other

5. What is your GPA? (Indicate “No GPA” if you do not have a GPA yet.)
   __________ GPA  __________ No GPA

6. What is your race?
   1 White/Caucasian  2 Black/African-American  3 Asian/Pacific
   4 Hispanic  5 Native American  6 Other
Appendix F

Almost Perfect Scale-Revised (APS-R), a Measure of Adaptive and Maladaptive Perfectionism

Please respond to the following items on a scale of 1-7, 1 meaning you “strongly disagree” with the statement and 7 meaning you “strongly agree” with the statement.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly agree</td>
</tr>
</tbody>
</table>

1. I often feel frustrated because I can’t meet my goals. (D)
2. I have high standards for my performance at work or at school. (HS)
3. My best just never seems to be good enough for me. (D)
4. I try to do my best at everything I do. (HS)
5. My performance rarely measures up to my standards. (D)
6. I think things should be put away in their place. (O)
7. I have high expectations for myself. (HS)
8. I am seldom able to meet my own high standards for performance. (D)
9. Doing my best never seems to be enough. (D)
10. I like to always be organized and disciplined. (O)
11. I am hardly ever satisfied with my performance. (D)
12. I set very high standards for myself. (HS)
13. I hardly ever feel that what I’ve done is good enough. (D)
14. If you don’t expect much out of yourself you will never succeed. (HS)
15. I am an orderly person. (O)
16. I rarely live up to my high standards. (D)
17. I expect the best from myself. (HS)
18. I often feel disappointment after completing a task because I know I could have done better. (D)
19. I often worry about not measuring up to my own expectations. (D)
20. Neatness is important to me. (O)
21. I am never satisfied with my accomplishments. (D)
22. I have a strong need to strive for excellence. (HS)
23. I am not satisfied even when I know I have done my best. (D)

Scoring Procedure: None of the items are reverse-coded. Scores are obtained by adding the reported scores according to the facets the individual items represent. “HS” signifies high standards, “D” signifies discrepancy, and “O” signifies order. Scoring high on “HS” and “O” means that the individual has high adaptive perfectionism whereas scoring high on “D” means the individual has high maladaptive perfectionism.
Appendix F, continued

Factor Loadings

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>APS11</td>
<td>.815</td>
<td></td>
</tr>
<tr>
<td>APS09</td>
<td>.806</td>
<td></td>
</tr>
<tr>
<td>APS03</td>
<td>.773</td>
<td></td>
</tr>
<tr>
<td>APS05</td>
<td>.769</td>
<td></td>
</tr>
<tr>
<td>APS13</td>
<td>.768</td>
<td></td>
</tr>
<tr>
<td>APS08</td>
<td>.726</td>
<td></td>
</tr>
<tr>
<td>APS18</td>
<td>.667</td>
<td></td>
</tr>
<tr>
<td>APS23</td>
<td>.631</td>
<td></td>
</tr>
<tr>
<td>APS16</td>
<td>.617</td>
<td></td>
</tr>
<tr>
<td>APS19</td>
<td>.606</td>
<td></td>
</tr>
<tr>
<td>APS21</td>
<td>.555</td>
<td></td>
</tr>
<tr>
<td>APS01</td>
<td>.451</td>
<td></td>
</tr>
<tr>
<td>APS02</td>
<td>.806</td>
<td></td>
</tr>
<tr>
<td>APS07</td>
<td>.800</td>
<td></td>
</tr>
<tr>
<td>APS17</td>
<td>.786</td>
<td></td>
</tr>
<tr>
<td>APS22</td>
<td>.778</td>
<td></td>
</tr>
<tr>
<td>APS12</td>
<td>.751</td>
<td></td>
</tr>
<tr>
<td>APS04</td>
<td>.679</td>
<td></td>
</tr>
<tr>
<td>APS10</td>
<td>.534</td>
<td></td>
</tr>
<tr>
<td>APS20</td>
<td>.529</td>
<td></td>
</tr>
<tr>
<td>APS15</td>
<td>.421</td>
<td></td>
</tr>
<tr>
<td>APS06</td>
<td>.350</td>
<td></td>
</tr>
<tr>
<td>APS14</td>
<td>.326</td>
<td></td>
</tr>
</tbody>
</table>
Appendix G

Life Orientation Test-Revised (LOT-R), a Measure of Generalized Optimism Versus Pessimism

Please respond to the following items on a scale of 1-7, 1 meaning you “strongly disagree” with the statement and 7 meaning you strongly “agree” with the statement.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly agree</td>
</tr>
</tbody>
</table>

1. In uncertain times, I usually expect the best.
2. It’s easy for me to relax. (Filler)
3. If something can go wrong for me, it will. \(^r\)
4. I’m always optimistic about my future.
5. I enjoy my friends a lot. (Filler)
6. It’s important for me to keep busy. (Filler)
7. I hardly ever expect things to go my way. \(^r\)
8. I don’t get upset too easily. (Filler)
9. I rarely count on good things happening to me. \(^r\)
10. Overall, I expect more good things to happen to me than bad.

*Scoring Procedure:* After reverse-coding the negatively worded items (as denoted by the superscript “\(^r\)”), sum the six non-filler items to produce the overall optimism score. Divide this score by 10 to obtain the score in the original metric. Higher scores signify greater optimism.
Appendix H

Meta-cognition

Please read each of the following statements. Rate the extent to which you agree or disagree with each statement. Even if you are unsure of an item, please answer it anyway.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly disagree</td>
<td>Strongly agree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. I thought a lot about my thoughts when working on the task.
2. I was aware of the way my mind worked when I was thinking through a task problem.
3. I monitored my thoughts while working on the task.
4. I rarely questioned my thoughts when working on the task.
5. I was constantly aware of my thinking when performing the task.
6. I paid close attention to the way my mind was working during the task.
7. I constantly examined my thoughts when working on the task.

Scoring Procedure: Reverse-score item 4. Sum scores on the items and divide by 7 to obtain score in original metric. Higher scores signify more meta-cognitive thoughts.
Appendix I

Emotion Control

Some difficulties may have arisen while working on the task(s). Please choose the response that best describes your reaction to these difficulties.

WHEN DIFFICULTIES AROSE…

1 2 3 4 5
Somewhat False Neither True Somewhat True False Or False True True

1. ...I did not allow myself to lose my composure.
2. ...I purposely continued to focus myself on the task.
3. ...I calmly considered how I could continue the task.
4. ...I allowed myself to be distracted by worrisome thoughts. r
5. ...I let myself become distracted. r
6. ...I let myself be sidetracked from the task. r
7. ...I was able to focus all my attention on the task.
8. ...I was able to motivate myself to continue.

Scoring Procedure: Items 4, 5, and 6 are reverse-scored. Sum up the responses to the 8 items and one obtains the amount of emotion control exhibited by participants. Higher scores signify higher emotion control.
Appendix J

Cognitive Appraisals

We are interested in your thoughts about the task. Please respond to the following items on a scale of 1-7, with “1” meaning “Strongly Disagree” and “7” meaning “Strongly Agree.”

<table>
<thead>
<tr>
<th>Item</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I view the task as a threat.</td>
<td></td>
</tr>
<tr>
<td>2. I expect the task to be stressful.</td>
<td></td>
</tr>
<tr>
<td>3. I often think about what it would be like if I did poorly on this task.</td>
<td></td>
</tr>
<tr>
<td>4. I believe the task could have negative consequences for me.</td>
<td></td>
</tr>
<tr>
<td>1. I view the task as a positive challenge.</td>
<td></td>
</tr>
<tr>
<td>2. The task provides me with an opportunity to meet a challenge.</td>
<td></td>
</tr>
<tr>
<td>3. I think the task represents a positive challenge to me.</td>
<td></td>
</tr>
<tr>
<td>4. I often think about what it would be like if I did well on the task.</td>
<td></td>
</tr>
<tr>
<td>5. I believe the task could have positive consequences for me.</td>
<td></td>
</tr>
</tbody>
</table>

**Threat construal**

**Challenge construal**

Scoring Procedure: Sum items together. Divide the threat total by 4 and the challenge total by 5 to obtain the score in its original metric. Higher scores reveal higher sense of experienced challenge and threat. None of the items require reverse scoring.
Appendix K

Self-Set Goals and Goal Commitment

Self-Set Goals

What is your goal for how many class schedules you can successfully create in the next trial?

_________________

Goal Commitment

We are now interested in your attitudes towards your goal for the next trial. Please respond to the following items on a scale of 1-5, with “1” signifying “strongly disagree” and “5” signifying “strongly agree.”

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly agree</td>
</tr>
</tbody>
</table>

1. It is hard to take this goal seriously. r
2. It is unrealistic for me to expect to reach this goal. r
3. It is quite likely that this goal may need to be revised, depending on how things go. r
4. Quite frankly, I don’t care if I achieve this goal or not. r
5. I am strongly committed to pursuing this goal.
6. It wouldn’t take much to make me abandon this goal. r
7. I think this goal is a good goal to shoot for.

Scoring Procedure: Items 1, 2, 3, 4, and 6 need to be reverse-scored so that high scores reflect high goal commitment. Once the negative items are reversed, sum up the responses to the 7 items to attain a measurement of goal commitment. Divide by 7 to attain score in original metric. Higher scores signify greater commitment to the goal.
Appendix L

Subjective Task Complexity

Please read each of the following statements. Rate the extent to which you agree or disagree with each statement.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not At all</td>
<td>Very</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. How difficult is performing this task?
2. How challenging is this task?
3. How discouraged do you feel when performing this task?
4. To what extent can you increase your performance by trying harder?
5. To what extent does this task require a lot of effort to be successful?
6. To what extent does this task require a lot of mental ability to be successful?
7. To what extent will you put forth a lot of effort to perform this task?

Scoring Procedure: Sum item responses and divide by 7 to attain score within the original metric. Higher scores signify greater perceived task complexity.
Appendix M

Task Satisfaction

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. How satisfied were you with your performance on the previous trial?
2. How satisfied will you be if you achieve the same performance level on the next trial?

*Scoring Procedure:* Sum the two items and divide by two to obtain score in its original metric. Higher scores signify greater satisfaction with the task.
Appendix N

Self-Efficacy Scale
Please read each of the following statements. Rate the extent to which you agree or disagree with each statement. Even if you are unsure of an item, please answer it anyway.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strongly disagree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Strongly agree</td>
</tr>
</tbody>
</table>

1. I have confidence in my ability to perform this task.
2. There are some activities required by this task that I cannot do well.\(^r\)
3. When my performance is poor, it is due to my lack of ability.\(^r\)
4. I doubt my ability to perform this task.\(^r\)
5. I have all the skills needed to perform this task very well.
6. Most students can do this task better than I can.
7. My future success on this task is limited due to my lack of skills.\(^r\)
8. I am very proud of my skills and abilities on this task.
9. I feel threatened when others watch me work.\(^r\)

Scoring Procedure: Items 2, 3, 4, and 5 need to be reverse-scored. Sum items together and then divide by 10 to obtain score in its original metric. Higher scores signify higher self-efficacy.
CONSENT TO PARTICIPATE IN RESEARCH
Department of Psychology
Wright State University
Dayton, OH 45435

Title of the Study: Performance on a Class Scheduling Task.

Purpose of Study: The investigators are interested how participants perform on a class scheduling task.

Activities: I understand that during the experimental session, I will be participating and creating class schedules on a computer. The task requires participants to form class schedules for mock students according to pre-specified guidelines and rules. Feedback will be given after each trial session regarding performance on the previous trials. I also understand that I will be completing various questionnaires throughout the experimental session.

Risks/Benefits: I understand that there is minimal risk associated with this experiment. The only potential risk is fatigue due to constant use of a computer. If I need a break, I understand that I will be able to take one without penalty. Although injury is highly unlikely, emergency medical assistance is available should the need arise. There are no direct benefits to participation other than contributing to the advancement of scientific understanding and becoming familiar with the procedures of psychological experimentation.

Confidentiality: All personal identifiers or information will be kept confidential. Individual information will not be identified in any published report.

Compensation: I understand that I will receive 1 extra credit point per half-hour of participation.

Freedom to Withdraw: I understand that research participation is completely voluntary and that I have the right to refuse to participate and withdraw from any part of the experiment at any point in time without penalty.

Availability of Results: I understand that I have access to a summary of results obtained from this experiment and can contact the principal investigator by June 2010. The results will not contain any personal identifiers.
Appendix O, continued

**Investigator Availability:** If I have questions or concerns regarding this experiment, I understand that I can contact the principal investigator (Zach Kalinoski, kalinoski.2@wright.edu) or the faculty advisor (Dr. Debra Steele-Johnson, debra.steele-johnson@wright.edu)

**Consent:** I understand that by signing my name below, I give my consent to participate in this research study.

**Signatures:**

<table>
<thead>
<tr>
<th>Signature of Participant</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

______________________________
Printed Name of Participant

______________________________
Zach Kalinoski
(Principal Investigator)
Date

______________________________
Debra Steele-Johnson, Ph.D.
(Faculty Advisor)
Date
Appendix P

Task Instructions

INTRODUCTION

The task you are about to work on involves creating class schedules for students. We are interested in collecting information about how people perform this task.

Throughout this task, you will use the LEFT ARROW, RIGHT ARROW, PAGE UP, PAGE DOWN, F1 to F7, and ENTER keys. These are all of the keys you will need to perform the task.

The purpose of the task is to create class schedules for students efficiently. As you work you will receive information about your performance.

Also, at various points during the task, we will ask you to answer some questions. Please read all the instructions very carefully. You will need this information to perform the task.

Be sure to read the instructions on each screen carefully BEFORE YOU PRESS THE SPACE BAR. Once you press the space bar to move to the next section, you CAN NOT return to a previous screen to reread any instructions.

TASK INSTRUCTIONS

We will now describe each part of the task and the task rules. Please read the following information carefully. You will need this information to perform the task.

To perform this task you must first learn what to do and how to perform actions using the keyboard. Please locate the printed keyboard diagram provided to you as we describe the necessary keys required for this task. The next section explains the purpose of each key.

The screen display is divided into 4 windows. Please locate the printed diagram provided to view an example of the window screen. The following keys will be used to perform actions in those windows:

LEFT ARROW: moves the cursor to the next window to the left, i.e., counterclockwise around the 4 windows
RIGHT ARROW: moves the cursor to the next window to the right, i.e., clockwise around the windows
UP ARROW: moves the cursor up within a window
DOWN ARROW: moves the cursor down within a window
PAGE UP: moves the cursor up to the previous course or previous student
PAGE DOWN: moves the cursor down to the next course or the next student
Appendix P, continued

F1 to F7: displays the task rules 1 through 7
ENTER KEY: completes the action you have taken

In this task you may perform 3 basic actions:
1) selecting courses from the course schedule
2) assigning courses to student schedules
3) reviewing previous student schedules.

The next screens (see diagram) describe the screen display and how to perform actions. The screen display is divided into 4 windows:
1) the course schedule window
2) the planning window
3) the review window
4) the information window.

The upper left quadrant of the window screen is an example of the course schedule window. This window displays the courses in the course schedule. You choose courses from this list to assign to students. This window displays one course at a time. For each course, you will see the course name and number. The next line shows any course prerequisites that might apply. The following lines display the course sections. Listed for each section are: course name, course number, section number, days of week, and time of day.

To select a course section from the schedule, you first use the RIGHT or LEFT ARROW key to move to the course schedule window. The window you are currently in is outlined by a light blue box. Once you are in the course schedule window, you select a course section by moving the selection box. The selection box is a small white rectangle that outlines a course.

There are a number of different courses you may choose from. You use the PAGE UP and PAGE DOWN keys to move from one course to another. You can see all of the courses in the schedule using the PAGE UP and PAGE DOWN keys. Also, there is often more than one section offered for a given course. You use the UP and DOWN ARROW keys to outline different sections of a course.

The upper right quadrant of the window screen is an example of the planning window. This window displays the student schedules to be completed and the course section assignments you have made. This window displays student schedules one at a time. For each student's schedule, you will see the student's social security number. The next line shows any special information pertaining to the student. The following lines display assigned course sections. Listed for each section are: course name and number, section number, days of week, and time of day.

You use the RIGHT or LEFT ARROW key to move to the course schedule window. The window will be outlined by a light blue box when you are working in it. Once you are in the
Appendix P, continued

planning window, you may select a space in a student's schedule by moving the selection box, the white rectangle.

You will work on a number of different student schedules. You use the PAGE UP and PAGE DOWN keys to move from one student’s schedule to the next student’s schedule. You use the UP and DOWN ARROW keys to outline different spaces in a student’s schedule. A selected course section will be assigned to the space outlined in a student’s schedule.

Once you have selected a course section and outlined a space in a student’s schedule, you can assign that course section by pressing the ENTER KEY. The course assignment will appear in the space you outlined in the student schedule. You may assign a course by pressing the ENTER KEY when you are in either the course schedule window or the planning window.

The lower right quadrant of the window screen is an example of the review window. This window displays any student’s schedule whether completed or not. The purpose of this window is to enable you to compare a student schedule you are working on in the planning window to any other student’s schedule. The review window displays student schedules one at a time. The same information is found in the planning window and the review window, including the student’s social security number, special information, course name and number, section number, days of week, and time of day.

You use the RIGHT or LEFT ARROW key to move to the review window. The window will be outlined by a light blue box when you are in it. Once you are in the review window, you may outline a space in a student’s schedule by moving the selection box, the white rectangle. You use the PAGE UP and PAGE DOWN keys to move from one student’s to another student’s schedule. You use the UP and DOWN ARROW keys to outline different spaces in a student’s schedule. You can look at, but not change, any information in this window.

The lower left quadrant of the window screen is an example of the information window. This window displays information about the task and about your performance. In the upper left corner of this window is a clock. The clock tells you how much time has gone by in the trial.

In the lower middle section of the window is “NO. SCH” which indicates the number of schedules you have completed. In the lower right section of the window is “NO. ERR” which indicates the number of errors you have made.

Also displayed in the information window are rule violation messages. If you break a rule, an error message will be displayed in the information window until the error is corrected.

Finally, the information window displays the task rules. You can review any rule by pressing the appropriate function key. F1 to F5 displays Rule 1 to Rule 5. When you press a function key, the rule will be displayed for 6 seconds or until you press a different function key.

Now, you are ready to begin the task.
## Appendix P, continued

### Course Schedule Window

<table>
<thead>
<tr>
<th>Chemistry 101</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chem 101 053010 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053011 M W F</strong> 9:00am - 10:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053012 T TH</strong> 8:30am - 10:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053013 M W F</strong> 12:00pm - 1:00pm</td>
</tr>
<tr>
<td><strong>Chem 101 053014 M W F</strong> 2:00pm - 3:00pm</td>
</tr>
<tr>
<td><strong>Chem 101 053015 T TH</strong> 11:00am - 12:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053016 T TH</strong> 2:00pm - 3:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053017 M W F</strong> 4:00pm - 5:00pm</td>
</tr>
<tr>
<td><strong>Chem 101 053018 M W F</strong> 6:00pm - 7:00pm</td>
</tr>
<tr>
<td><strong>Chem 101 053019 T TH</strong> 5:00pm - 6:30pm</td>
</tr>
</tbody>
</table>

### Planning Window

<table>
<thead>
<tr>
<th>ST 822631340</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chem 101 053010 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053011 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053012 T TH</strong> 8:30am - 10:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053013 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053014 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053015 T TH</strong> 11:00am - 12:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053016 T TH</strong> 2:00pm - 3:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053017 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053018 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053019 T TH</strong> 8:00am - 9:00am</td>
</tr>
</tbody>
</table>

### Information Window

<table>
<thead>
<tr>
<th>TIME 1:12 POSSIBLE 8 GOAL 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>One section of each class per plan</td>
</tr>
<tr>
<td>Cannot have overlapping course times</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CUM -SCH</th>
<th>NO. SCH</th>
<th>NO. ERR</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CUM</th>
<th>SCH 20</th>
<th>ERR 100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Review Window

<table>
<thead>
<tr>
<th>ST 822631340</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chem 101 053010 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053011 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053012 T TH</strong> 8:30am - 10:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053013 M W F</strong> 8:30am - 10:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053014 M W F</strong> 8:30am - 10:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053015 T TH</strong> 11:00am - 12:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053016 T TH</strong> 2:00pm - 3:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053017 M W F</strong> 5:00pm - 6:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053018 M W F</strong> 5:00pm - 6:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053019 T TH</strong> 5:00pm - 6:30pm</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ST 822631340</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chem 101 053010 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053011 M W F</strong> 8:00am - 9:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053012 T TH</strong> 8:30am - 10:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053013 M W F</strong> 8:30am - 10:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053014 M W F</strong> 8:30am - 10:00am</td>
</tr>
<tr>
<td><strong>Chem 101 053015 T TH</strong> 11:00am - 12:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053016 T TH</strong> 2:00pm - 3:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053017 M W F</strong> 5:00pm - 6:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053018 M W F</strong> 5:00pm - 6:30pm</td>
</tr>
<tr>
<td><strong>Chem 101 053019 T TH</strong> 5:00pm - 6:30pm</td>
</tr>
</tbody>
</table>
Appendix Q

Feedback Cuing Screens

EMT

Pay special attention to the feedback given to you on the lower left screen during the task. If you make an error, an error message will appear to allow you to better understand and correct the errors.

EXP

Feel free to explore all aspects of the screen for information that will be useful for your task performance.

EAT

You will be able to gauge your performance during the task by looking at your cumulative score (denoted as “CUM”). This score factors in number of schedules (20 points each) completed and number of errors committed (5 points each). This score will appear in the lower left screen.
Appendix R

Debriefing Form

The experiment you just completed sought to understand the influences of different training systems on one’s performance. We were also interested in how individual perceptions and stress appraisals explain the differential effects of the training interventions on one’s performance in the class scheduling task. Finally, we also wanted to know whether or not personality affected this relationship.

Specifically, some of the training literature is suggesting that making errors during the training session can have beneficial effects on future performance on similar and different activities. Previous research suggested that thinking about one’s thoughts and controlling negative emotions that result from making many errors during training explains the beneficial effects of errors on future performance. We were interested in replicating these effects as well as looking at an additional explanation. Past research has also been inconsistent in regards to personality and its interactive effects on the training-performance relationship.

It is vital that you do not discuss the details of the training manipulations or the class scheduling task so as to protect the integrity of the research. Disregarding this warning would jeopardize the hard work and time that previous participants put in to contributing to this area of research and scientific advancement. Advanced knowledge of the experiment can negatively affect the results obtained.

Thank you very much for your participation in this research. Scientific progress in human effectiveness could not be done without your help and assistance. If you have any questions or comments, please contact the principal investigator, Zach Kalinoski (kalinoski.2@wright.edu) or Dr. Debra Steele-Johnson (debra.steele-johnson@wright.edu).