MCCONNELL, MELISSA D., Ph.D. Examining the Benefits of Feedback: Are Monitoring Skills Implicated in Successful Performance? (2009)
Directed by Dr. Stuart Marcovitch. 142 pp.

The current set of experiments was designed to investigate whether monitoring skills are implicated in the effectiveness of receiving feedback on performance. This was examined by determining whether receiving feedback improves the retention of correct responses as well as improves memory performance (Experiment 1), whether participants detect and use false feedback (Experiment 2), and whether young children's memory performance improves from receiving feedback (Experiment 3). In addition, confidence ratings were taken as a measure of participants' ability to monitor their performance. The results revealed participants' confidence in their original responses influenced the effectiveness of feedback, and participants used the feedback to influence their memory and monitoring performance. These results imply participants' confidence influences the effects of receiving feedback.

EXAMINING THE BENEFITS OF FEEDBACK: ARE MONITORING SKILLS IMPLICATED IN SUCCESSFUL PERFORMANCE?

by

Melissa D. McConnell

A Dissertation Submitted to
The Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro

2009

Approved by

_____
Committee Co-Chair

_____
Committee Co-Chair

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Co-Chair _____

Committee Co-Chair _____

Committee Members _____

_____

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Feedback is commonly used to enhance classroom learning. Because of its wide use and promising applications in education, researchers have been interested for years in understanding what factors influence the effectiveness of feedback and how feedback improves memory performance (e.g., Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Kulhavy, 1977; Thorndike, 1913). Providing feedback after classroom assessment is a common technique used by teachers to help students learn material and improve performance on later tests (e.g., Bangert-Drowns et al. 1991), which is why researchers in the field of psychology and education are interested in investigating the possible benefits of receiving feedback. Although most research is conducted in laboratory settings (e.g., Anderson, Kulhavy, & Andre, 1971, 1972; Andre & Theiman, 1988; Kulhavy & Anderson, 1972), at times researchers have also conducted studies incorporated in the existing classroom lesson plans (e.g., Brosvic, Epstein, Dihoff, & Cook, 2006; Dihoff, Brosvic, & Epstein, 2003; Dihoff, Brosvic, Epstein, & Cook, 2004).

As defined by Kulhavy (1977), feedback is a procedure used to inform a learner of the correctness of a response and it is distinguished from reinforcement, which increases the probability of a response. Consequently, feedback's role in performance is not to *increase* a particular response but to *offer information* about a particular memory response. Feedback can contain two components: verification and elaboration (Kulhavy,

1977; Stock & Kulhavy, 1989). Verification informs the learners of the correctness of their response, while elaboration indicates information about the response. The most common type of elaboration is providing the correct alternative (Anderson et al., 1972; Kulhavy, 1977; Pashler, Cepeda, Wixted, & Rohrer, 2005; Stock & Kulhavy, 1989). In academic feedback paradigms, researchers typically provide participants with feedback about the correctness of their original response and test whether the feedback improved their memory performance.

Academic feedback paradigms consist of: (a) an initial test (Test 1), (b) feedback session, and (c) a final test (Test 2). Furthermore, the information being tested can either be learned in a study period, or the exam can test general knowledge information learned outside of the experiment. In many original studies testing feedback, the experimenters provided a study period using a technique termed programmed instruction. In this method, participants viewed one short paragraph and answered one or two test questions before reading a new paragraph (e.g., Anderson, Kulhavy, & Andre, 1972). However, the study periods in contemporary experimental techniques are designed to match more closely with classroom learning situations. In these studies participants may read a journal article, study a foreign language, or read a textbook chapter (e.g., Kang, McDermott, & Roediger, 2007; Pashler et al., 2005). The most common variable assessed is the proportion of Test 2 questions answered correctly (i.e., Test 2 proportion correct). Examining Test 2 proportion correct after receiving feedback demonstrates whether feedback improves overall performance. Participants can improve their overall performance by correcting their errors from Test 1 or retaining their correct responses. Error correction is measured as the proportion of

errors on Test 1 that are corrected on Test 2, while the retention of correct responses is measured as the proportion of items answered correctly on both tests. Many factors, such as the timing and type of feedback, influence whether receiving feedback improves participants' memory performance (e.g., Phye, 1979; Phye & Andre, 1989; Sturges, 1972).

Typical Effects of Receiving Feedback on Memory Performance

McConnell (2007) examined 52 classroom and lab experiments investigating the effectiveness of feedback on memory performance and reported that the most commonly reported measure is Test 2 proportion correct. In 88% of the experiments that reported Test 2 proportion correct, performance improved when feedback was provided (e.g., Clariana, Ross, & Morrison, 1991; Kulhavy & Anderson, 1972). Examining Test 1 correct and incorrect responses separately determines how feedback improved performance. Stock and Kulhavy (1989) proposed that feedback works differently for originally correct and incorrect responses (see also Pashler et al., 2005). They suggested that feedback provided after a correct response strengthens and increases the durability of the response over time. In contrast, feedback after an error needs to: (a) eliminate the incorrect response, (b) substitute the correct response, and (c) strengthen the durability of the correct response over time. When examining initial errors and correct responses separately, McConnell (2007) found that receiving feedback led to increased error correction in 95% of the experiments reporting error correction (e.g., Kulhavy, Yekovich, & Dyer, 1976, 1979; McConnell & Hunt, 2007; Pashler et al., 2005). On the other hand, feedback had a much less consistent effect on the retention of correct responses. Only 50% of the experiments that reported correct retention found that receiving feedback

increased the retention of correct responses (e.g., Dihoff et al., 2003; Dihoff et al., 2004; Kulhavy et al., 1976, McConnell & Hunt, 2007, Exp. 1).

The Effect of Confidence and Feedback on Improving Memory Performance

This clearer effect of feedback on error correction led Kulhavy and colleagues (e.g., Kulhavy, 1977; Kulhavy et al. 1976, 1979; Kulhavy, Stock, Hancock, Swidell, & Hammrich, 1990; Stock & Kulhavy, 1989) to propose that the main goal of feedback is to correct metacognitive errors. Metacognition is typically defined as knowledge of one's cognitive processes (Hackler, 1998; Matlin, 2005; Nelson & Narens, 1990). This typically involves monitoring ones learning and thoughts and then using the monitoring to control or regulate behavior (e.g., Flavell, 1979; Hacker, 1998; Koriat & Goldsmith, 1996; Nelson & Narens, 1990). Monitoring accuracy is typically measured by participants predicting how likely they will remember the item on a later test (Connor, Dunlosky, & Hertzog, 1997; Dunlosky, Rawson, & Middleton, 2005; Koriat & Goldsmith, 1996; Nelson & Narens, 1990), or by participants providing their confidence in their response, (e.g., Bisanz, Vesonder, & Voss, 1978; Koriat & Goldsmith, 1996). The findings from research demonstrate that although participants may be overconfident in their monitoring (e.g., Koriat, Lichtenstein, & Fischhoff, 1980) or underconfident (e.g., Serra & Dunlosky, 2005), they do monitor and regulate their performance successfully (e.g., Dunlosky & Thiede, 1998; Koriat & Goldsmith, 1996; Nelson & Narens, 1990; Son & Metcalfe, 2000). Participants tend to provide higher levels of confidence for questions answered correctly and lower levels of confidence for those answered incorrectly (e.g., Bisanz et al., 1978; Koriat & Goldsmith, 1996).

Kulhavy and Stock (1989) proposed a relationship between learners' confidence judgments and the effectiveness of feedback, and emphasized the importance of factoring in participants' confidence when examining the effectiveness of feedback. Borrowing heavily from metacognition research, they propose that participants set an expectation for performance (typically the goal is correct responding), and assess how much effort will be required to meet their goals. If their perception of the task demands does not match their current state of knowledge, they will exert more effort to meet the goals (see also Nelson & Narens, 1990). When the learners assess their confidence in their responses, the ratings reflect whether they believe they met their learning goals. Researchers investigating metacognition typically examine participants' ability to monitor their performance by observing their confidence judgments, and suggest that several factors influence participants' confidence judgments and the belief that they met the learning goals: how much of the information they understand (e.g., Tiede & Leboe, 2009), how well they generally perform on the task (Koriat, 2008), familiarity with the topic (e.g., Metcalfe, 1993), and how much and how quickly information came to mind (e.g., Koriat, 1993).

Confidence and Error Correction

Kulhavy and colleagues (e.g., Kulhavy, 1977; Kulhavy et al., 1976, 1979; Kulhavy et al., 1990; Stock & Kulhavy, 1989) also stressed that the effects of receiving feedback may differ as a function of participants' confidence in their initial responses. Without receiving feedback there is a strong relationship between initial confidence and correct responding on Test 2 (e.g., Kulhavy, et al., 1976, 1979). Participants are more

likely to respond on with Test 2 with a highly confident response from Test 1, regardless if correct, than with a low confident response. A different pattern emerges, however, after feedback is provided. Kulhavy and Stock (1989) proposed that feedback is compared to the original response, and importantly, participants evaluate the accuracy of their confidence in that original response. If there is a discrepancy between what the participants believed they knew and what they actually know, behavior is modified to reduce the discrepancy. Interestingly, a larger discrepancy is detected when participants were highly confident in their errors perhaps prompting them to exert more effort to understand and remember the feedback. Kulhavy et al. (1976, 1979) had participants learn information and then take a multiple-choice test covering the studied material. They provided participants with feedback after each test question, and participants controlled the amount of time they studied the feedback. Kulhavy et al. found participants spent more time studying the feedback after committing errors in high confidence than in low confidence. Kulhavy et al. (1990) allowed participants the opportunity to restudy the text after receiving feedback, and they found that participants were more likely to restudy the text after receiving feedback on high confident errors. This increased effort resulted in participants correcting more errors committed in high confidence than in low confidence.

Butterfield and Metcalfe (2001, 2006) demonstrated that the relationship between confidence in Test 1 responses and feedback persists even when participants do not have control of the amount of time they study the feedback. They proposed that participants provide more attention to high confident errors than to low confident errors. Butterfield and Metcalfe (2006) gave participants 619 general knowledge questions that all required

a single word answer. Participants answered the questions, gave a confidence rating, and received feedback immediately after the response. The feedback provided appeared in green if participants answered correctly and red if participants answered incorrectly. To examine whether participants paid more attention to surprising feedback, participants heard a tone randomly throughout the experiment, and were asked to hit the spacebar after hearing it. Participants were less likely to press the spacebar when viewing feedback that indicated their responses made in high confidence were incorrect compared to when the feedback indicated their responses made in low confidence were incorrect. This implies participants were paying more attention to the surprising feedback. Butterfield and Metcalfe used the term *hypercorrection with feedback effect* to describe the finding that participants corrected more errors committed in high confidence than in low confidence.

Confidence and Correct Retention

Kulhavy and colleagues (e.g., Kulhavy, 1977; Kulhavy et al., 1976, 1979; Kulhavy et al., 1990; Stock & Kulhavy, 1989) emphasized the importance of confidence when examining error correction, but not when examining the retention of correct responses. They did report that participants were least likely to study feedback after a highly confident correct response, and noted that participants studied the feedback for low confident errors and correct responses only briefly. They suggested that low confidence reflects insufficient knowledge of the material, but did not speculate on the role confidence may play. Butler, Karpicke, and Roediger (2008), however, recognized the potential role that confidence in original responses has in the retention of correct

responses. Butler et al. stressed that low confident correct responses also indicate a metacognitive error in monitoring, and that receiving feedback reduces the discrepancy by increasing confidence in these responses. They suggested that participants may have low confidence in a correct response for different reasons: (a) they made a lucky guess, (b) they have only some familiarity with the question, or (c) they were deciding between two options. In these situations, receiving feedback confirming the responses should increase the likelihood of retaining the responses. Participants answered questions on a multiple-choice test, provided confidence judgments, received feedback, and answered questions on a cued-recall test. Participants were forced to respond on the multiple-choice test, but were allowed to refrain from responding on the cued-recall test. They found that providing feedback on the multiple-choice test enhanced the retention of correct responses on the cued-recall test. Butler et al. proposed that participants answered some questions correctly on the Test 1 recognition that were weak (or only slightly familiar), and receiving feedback strengthened those responses. These correct responses would probably not have been recalled if the initial test was the cued-recall test. Because participants received feedback on low confident correct responses on a multiple-choice test, they were now more willing to report those answers on the cued-recall test. Butler et al. concluded that feedback enhanced the retention of correct responses by making participants more likely to report an answer, and that receiving feedback enhances correct retention when using special testing formats in which participants are given a recognition test first and a recall test later.

By emphasizing the importance of testing format when examining the retention of correct responses, Butler et al. (2008) found that receiving feedback enhanced the retention of correct responses from an initial recognition test onto a later cued-recall test. Based on the suggestion that feedback increased the retention of low confident correct responses, an effect of feedback on correct retention should be apparent when using methodology that encourages enough low confident correct responses, even when not using the testing formats emphasized by Butler et al. Further research is required to determine whether feedback increases the retention of correct responses when participants answer enough questions correctly in low confidence, and not simply when a particular testing format is used (recognition and then recall). If feedback increases the retention of correct responses under a wider range of conditions, than educators and researchers may need to emphasize the importance of receiving feedback for error correction *and* enhancing correct retention.

Confidence in the Feedback

A lot of emphasis has been placed on the importance of participants' confidence in their original responses. But if Kulhavy and colleagues are correct (e.g., Kulhavy, 1977; Kulhavy et al., 1976, 1979; Kulhavy et al., 1990; Stock & Kulhavy, 1989), and the function of feedback is to correct metacognitive errors, than participants' confidence in the feedback may also be important. As Kulhavy and Stock (1989) noted, when feedback indicates a discrepancy between the learners' knowledge and their goal for learning, participants change their behavior to reduce the discrepancy (e.g., pay more attention to the feedback or study the feedback longer). However, if feedback provides incorrect

information, it may falsely indicate a discrepancy between the learners' goals and knowledge. Responding to the feedback may depend on whether participants notice that the feedback provides incorrect information, and whether they can ignore it. If participants have reason to doubt the information provided in the feedback they may be less likely to use the information, and may exert added effort to remember their original responses.

One method of manipulating confidence in the feedback is to alert the participants that some or all of the feedback is false. If participants were given an explicit reason to doubt the feedback, then they may be able to ignore it and respond with their original answers. Stock, Kulhavy, and Pridemore (1992) touched on the potential importance of participants' confidence in the feedback by warning participants that some of the feedback they were to receive may provide incorrect information about the correctness of their responses. Stock et al. provided true and false verification feedback, and found that when participants received feedback that falsely confirmed their response was right, they studied it longer than when the feedback falsely indicated their response was wrong. In addition, when the true feedback correctly confirmed their response was right, they studied it longer than when the feedback correctly indicated that their response was wrong. This indicates participants behaved similarly after receiving true and false feedback. At first glance, it may appear contradictory that the participants spent more time reading the feedback that confirmed a response. However, because the feedback only indicated whether the response was correct or incorrect, studying the feedback that indicated an incorrect response was not helpful (there was no correct alternative to study).

Participants could only improve their performance by reviewing the feedback that indicated their response was correct. Because Stock et al. did not provide the participants with an alternative response, it remains to be seen whether participants would use false feedback when provided with more information.

No other studies using an academic feedback paradigm provided warnings that some of the feedback may provide incorrect information, but several studies in similar paradigms provided false feedback and a warning. In one of the paradigms, the hindsight bias paradigm, the influence of receiving feedback on memory for the original response is examined to determine whether participants automatically integrate feedback into their existing knowledge. Research investigating the hindsight bias shares many similar features to feedback studies, and occurs when participants make a judgment, are told the solution, and then misremember their original judgment as closer to the solution than the original response. Most research demonstrating hindsight bias suggests that when participants receive feedback it interferes with their memory for their original responses, and they automatically integrate the feedback into their existing knowledge (Fischhoff, 1975, 1977; Fischhoff & Beyth, 1975; Pohl & Hell, 1996). In a typical hindsight bias study, participants make a judgment, receive feedback about their judgment, and are asked to remember their original judgment. Several studies found that participants' estimates of their original judgment shifts closer to the feedback (Fischhoff, 1975, 1977; Fischhoff & Beyth, 1975; Hasher, Attig, & Alba, 1981; Pohl & Hell, 1996). Although these findings suggest participants automatically use the feedback, if participants have a reason to doubt the feedback, they may be less likely to integrate the feedback into their

current knowledge. Hasher et al. (1981) investigated this by warning participants that some of the feedback was false. They asked participants general knowledge questions, and participants answered the questions and gave their confidence. After providing the answers, participants received feedback about the responses, and were then warned explicitly that some of the feedback was false. At test participants were told to answer with their original responses and confidence, and participants who received the warning gave responses and confidence ratings more similar to their original responses than the feedback. This confirms that if participants are given an explicit reason to doubt information, they are able to ignore it and respond with their original memories (e.g., Erdfelder & Buchner, 1998).

Although participants can sometimes ignore false feedback when provided with a warning, it is unknown whether participants can detect false feedback without being warned. This is important because individuals receive information frequently, but they do not typically have someone indicating whether the information is correct. Individuals have to judge for themselves the correctness of the information and decide whether to use it. Furthermore, determining whether participants can detect false feedback can speak to the ability to monitor incoming information. In a study by Hanaba and Zaragoza (2006), participants were forced to respond to questions about an event they witnessed. After answering some questions, participants received confirmatory feedback ("Yes that is correct") on questions they answered incorrectly. One week after the test, participants completed a recognition test, and participants were more likely to answer these questions again incorrectly after receiving false feedback confirming their errors. Hanaba and

Zaragoza's study does not completely address whether participants detect and use false feedback because the false feedback was always in the form of confirming an incorrect answer. Researchers have not yet investigated providing misinformation for responses participants answered correctly. This is a critical manipulation to test whether participants can detect the correctness of the feedback. When participants receive false feedback that confirms an incorrect response, the feedback matches what they already believed was the correct response. Yet, when participants receive feedback that refutes a correct response, the feedback may provide information that conflicts with their current knowledge. Only in the situation where the feedback conflicts with current beliefs would there be a reason to question the feedback.

Another area that sheds light on participants' ability to detect false feedback is the misinformation effect paradigm. In a typical misinformation effect task, participants view an event and then receive a narrative that contains true information but also critical misinformation about the original event. Participants are then tested and asked to answer questions about the original event. Participants' performance is impaired by receiving the misinformation (e.g., Lindsay & Johnson, 1989a; Loftus, 1979; Loftus, Miller, & Burns, 1978; McCloskey & Zaragoza, 1985). For example, Loftus et al. presented participants with slides of an accident and later heard a narrative describing the accident. The narrative contained some information that conflicted with the original scene depicted in the slides. Participants then answered questions about the slides, but they answered questions incorrectly because of the misinformation presented in the narrative. Studies providing misinformation consistently find that participants incorporate the

misinformation and will incorrectly respond with the false information from the narrative (e.g., Lindsay & Johnson, 1989a; Loftus et al., 1978; Payne, Toglia, & Anastasi, 1989). These findings suggest participants may not detect the false feedback in an academic feedback paradigm. However, there is an important difference between a typical misinformation effect study and a typical feedback study. Namely, the feedback about the original response comes after participants have responded on Test 1 in a feedback study, but before participants have responded about the original event in a misinformation study. Participants may be less likely to be influenced by false feedback if they have already responded to an event. Previous findings suggest participants' memories are improved by retrieving information (e.g. Bjork, 1988; Roediger & Karpicke, 2007; Tulving, 1967), and these memories may be less susceptible to false information.

Feedback and Preschool Children

Although many researchers demonstrate that college students benefit from receiving feedback (e.g., Anderson et al., 1972; Brosvic, et al., 2006; Kulhavy & Anderson, 1972; Phye & Baller, 1970), there is less work with young children receiving feedback. No researchers have directly examined whether preschool participants benefit similarly from receiving feedback as do college students, and this is an important age to assess the skill as these children are transitioning to formal schooling. They will be expected to learn new information in a structured manner at a rapid pace, and educators and parents rely on feedback to maximize young children's ability to learn new information. It is important to determine to what extent preschool children's learning abilities may benefit from receiving feedback. Only a few studies have even assessed the role of feedback in

participants younger than the college age range: three studies tested high school and middle school participants (Clariana et al., 1991; Kulhavy & Anderson, 1972; Kulhavy, et al., 1990) and three used participants in elementary school (Markowitz & Renner, 1966; Peeck & Tillema, 1978-1979; Peeck et al., 1985), and in most studies, feedback led to greater Test 2 performance. For example, Peeck et al. (1985) examined the effect of feedback on children's memory performance. In their study, fifth grade children read a text, answered multiple-choice questions, received feedback, and then answered the questions again. Children who received feedback corrected more errors than children who did not receive feedback.

Preschool is an interesting age group to investigate the benefits of feedback because young children appear to monitor less well than do adults (Bisanz et al., 1978; Flavell, Friedrichs, & Hoyt, 1970; Pressley, Levin, Ghatala, & Ahmad, 1987). Because one of the goals of feedback is to inform the learner of a metacognitive error, if young children have difficulties monitoring, they may not benefit from receiving feedback as do older children with more monitoring experience. Flavell et al. (1970) had 5-year-old and 8-year-old children study items as long as needed to obtain perfect recall. The 5-year-olds did not obtain perfect recall, despite claiming they studied the information adequately whereas, the older children had perfect or near perfect recall. In another experiment, Pressley et al. (1987) had 6- to 11-year-old children predict their performance after taking a multiple-choice task and found that younger children had poorer monitoring. Furthermore, Bisanz et al. (1978) found that 6-year-old children had more difficulty monitoring the accuracy of their responses when discriminating correct and incorrect

answers. These examples of monitoring difficulties with 5- and 6-year old children suggest they may not benefit as well from receiving feedback as college students.

There also may be other reasons to suspect that young children do not benefit from receiving feedback. First, according to an information processing framework of development, children may not be capable of benefiting from strategies that help college students. For example, Miller, Seier, Barron, and Probert (1994) asked preschool children to find and remember where objects were located, and they found children who used the most effective strategy to remember the objects did not actually benefit from using the strategy. Bjorklund and Harnishfeger (1987) found that providing children with strategies actually hurt their performance because using the strategies took away from other cognitive resources needed to complete the task. It may be that receiving feedback and trying to remember the feedback causes children more difficulty in processing the question, remembering the stories, or remembering their original responses.

Second, preschool children are prone to perseverative errors. A perseverative error occurs when children respond with correct responses from earlier trials that are now errors on the current trials. For example, Cinan, Fakultesi, and Bolumu (2006) gave 6- to 9-year-old children the Wisconsin card sorting task (Grant & Berg, 1948), in which children sorted cards that varied on three dimensions (i.e., shape, color, and number). The children were not told which dimension they should sort by for each trial, but they were given feedback after each trial. After several trials the sorting dimension changed, and a perseverative error occurred when children persisted on sorting by an old dimension even after receiving feedback that they sorted incorrectly. Cinan et al. (2006) found that the 6-

and 7-year-old children made more perseverative errors than did the 8- and 9-year-old children, consistent with a robust literature that suggests young children are prone to perseverative errors (e.g., Crone, Ridderinkhof, Worm, Somsen, & van der Molen, 2004; Crone, Somsen, Zanolie, & van der Molen, 2006; Zelazo, Muller, Frye, & Marcovitch, 2003). Zelazo, Frye, and Rapus's (1996) findings demonstrated the persistence of perseveration errors as they found preschool children continued to make perseverative errors on a card sorting task, even when they expressed knowledge of the correct rule. Because preschool children are prone to perseverative errors, receiving feedback may not be enough information to help them overcome them. As mentioned previously, the goals of feedback provided after an error are to substitute the error, replace the error, and strengthen the correct response. Because children are more likely to perseverate, it may be more difficult for them to inhibit and replace those errors after receiving feedback.

Memory Explanations

Original explanations of the misinformation effect described the effect as occurring through automatic updating of the original memory trace. As described in many studies, memory for the misleading information overwrites the original memory and forms a new integrated trace (e.g., Loftus 1979, 1992, 2003, 2005; Loftus & Loftus, 1980; Loftus et al., 1978; Payne et al., 1989). Other researchers suggest the original memory and the misinformation form separate traces; they explain the misinformation effect as caused from problems with accessibility or interference, (e.g., Bekerian & Bowers, 1983; Christiaansen, & Ochalek, 1983), source monitoring, (e.g., Abeles & Morton, 1989; Lindsay, 1993; Lindsay & Johnson, 1989a, 1989b; Rantzen & Markham,

1992; Templeton & Wilcox, 2000), and bias (e.g., Bekerian & Bowers, 1983; McCloskey & Zaragoza, 1985a, 1985b; Payne et al., 1989; Templeton & Wilcox, 2000).The misinformation effect is often explained using a source monitoring framework.

McConnell (2007) further speculated that receiving feedback could be viewed usefully from a source monitoring framework defined as a "set of processes involved in making attributions about the origins of memories, knowledge, and beliefs" (Johnson, Hashtroudi, & Lindsay, 1993, p. 3). Memory characteristics, such as perceptual, semantic, temporal, and cognitive operations, are incidentally formed during the original encoding and may provide information about the source of the memory. For example, the meaning of the information, the context the information occurred in, and any physical features of the information are all formed when the memory is encoded. Memories are not directly tagged or labeled with the source, but instead the memory characteristics are evaluated and a source is attributed through a decision process. During remembering, the information is judged based on the quantity and quality of the memory detail and then compared to what is expected from memories from different sources. If the memory has few perceptual details, it is judged to be from a source that provides few perceptual details (e.g., a dream). Often source monitoring occurs automatically, but difficult situations require a more systematic analysis of the memory qualities (Johnson, 2006; Johnson et al., 1993).

According to a source monitoring explanation, the feedback and the original response should be viewed as two distinct traces, or sources of information. At Test 2, information comes to mind and a participant may attribute correctly the sources of competing information.

Manipulations that reduce the quality of the memory characteristics and reduce the quality of the decision processes weaken source monitoring. For example, divided attention during encoding or source attribution worsens source monitoring judgments (Troyer, Winocur, Craik, & Moscovitch, 1999). If source monitoring is a primary influence on feedback, then manipulations that affect source monitoring should similarly affect the ability to discriminate between the original response and feedback. Johnson et al. (1993) argued that external source monitoring requires discerning two externally generated sources. Viewing feedback as a second source of information is a form of external source monitoring. The participants must discriminate between two external sources (Test 1 response and feedback).

Failure to discriminate source need not necessarily worsen performance. If, at test, participants remember a response but cannot discriminate the source, they may still answer correctly—but only in certain circumstances. Participants will answer correctly, without discriminating source, if an original correct response or the feedback comes to mind. According to a source monitoring framework, participants form memory records when completing the original test and while receiving feedback. These records include any perceptual details of the item, what they were thinking, and any cognitive operations they performed when forming the memory record. If participants are trying to discriminate the source of the memories when completing Test 2, they will be judging the memories that come to mind while answering the questions. When a memory comes to mind they will evaluate the record of the memory and judge the probability it was formed when receiving feedback or completing Test 1 (Johnson et al., 1993). For example, if participants realized they chose the incorrect response at the time of

the feedback, then these thoughts are part of the memory record for the feedback, which may be particularly helpful in attributing source.

Findings from a number of other feedback studies are consistent with a source monitoring framework. First, several studies find that participants can still recognize their original responses after receiving feedback (e.g., Brosvic et al., 2006; Dihoff et al., 2004; Peeck & Tillema, 1978-1979). From a one-trace theory perspective, the feedback should impair memory of the original response and would not predict participants should be able to remember their original responses. A source monitoring framework explains how participants who receive feedback are still able to remember their original responses because the feedback exists as another trace (or source of information). Second, participants who are able to recall or recognize their original responses actually perform better on Test 2 than participants who are unable to recall or recognize their original responses (e.g., Butterfield & Metcalfe, 2001; Peeck et al., 1985). Peeck et al. provided feedback after errors on a multiple-choice test and asked participants to respond with their original response and the correct response on a later test. They found that participants were more likely to answer Test 2 questions correctly when they correctly identified their original errors. Butterfield and Metcalfe asked participants to answer general knowledge questions, and participants received feedback on their performance. After a brief delay they were asked to respond with the three responses that came to mind, and then identify the correct response. They found that participants were more likely to answer Test 2 questions correctly if they included their original error in the three responses that came to mind. At first glance it is surprising that participants would be

more likely to answer questions correctly on Test 2 if they still remember their original incorrect response, but it may be that participants who remember their original response during Test 2 recognized that the feedback indicated a metacogntive error and spent more time reflecting on the feedback.

The third piece of evidence to support a source-monitoring framework is the superior performance after delayed as opposed to immediate feedback. According to a source monitoring framework, as two sources occur farther apart in time their similarity should lessen because the context may change, resulting in improved source monitoring (e.g., Drummey & Newcombe, 2002; Johnson et al., 1993). Items occurring further apart in time should have different contexts and may be easier to discriminate because of different temporal cues (Johnson et al., 1993). Therefore, as the delay between the original responses and the feedback increases, the context cues should increase and discrimination becomes easier. This increase in source monitoring may result in an increase in Test 2 performance with delayed feedback. Phye and Baller's (1970) findings support this idea. In their experiment, students studied information and received Test 1 feedback after a short or long delay. They found that participants answered a greater percentage of Test 2 items correctly when receiving feedback after a longer delay (see Hunt, Smith, & McConnell, 2009; Peeck & Tillema, 1978-1979; Phye, 1979; Phye & Andre, 1989).

Young children also demonstrate successful source monitoring in some contexts, suggesting this same framework may be appropriate for explaining how feedback improves their memory performance. Foley, Johnson, and Raye (1983) demonstrated that

6-year-old children discriminated external sources of information to the same degree as young adults. The participants watched two experimenters read facts of information, and the 6-year-olds successfully discriminated which experimenter read which fact.

Templeton and Wilcox (2000) proposed a developmental framework for children's ability to handle multiple sources of information. In their study, children viewed a videotape of a crime, and then listened to a narrative about the crime. The narrative contained both true and false information regarding the original crime. The children also participated in a false belief task (see Wimmer & Perner, 1983) where participants were asked what was in a box of crayons, and then after being shown it contained candy they were asked what their friend would believe was in the box. Templeton and Wilcox found that participants who failed the false belief task (answered that their friend would believe candy was in the box) were also more likely to suffer from the misinformation in the narrative. They concluded that children who failed the false belief task could only hold one representation in mind, and because the misleading narrative provided information more recently that was the only representation they held. These findings suggest in an academic feedback paradigm younger and older children may perform well for different reasons. Younger children may benefit from receiving feedback because they only remember the feedback, while older children, with successful source monitoring skills, may perform well because they can distinguish between their original response and the feedback.

The Effects of Feedback on Monitoring Performance

Educators are interested in providing feedback because of known examples of it improving memory performance on a later test, but receiving feedback may also improve test scores because it may improve students' ability monitor their performance. Koriat and Goldsmith (1996) found that participants monitor accurately in some situations but have great difficulty in other situations. In their first experiment, participants correctly monitored their performance as they answered general knowledge questions and provided confidence judgments by giving an estimate of how likely their answer was to be correct. In the second experiment deceptive general knowledge questions were added. In the deceptive questions, one of the incorrect alternatives appeared more likely to be correct than the actual correct response (e.g., What is the capitol of Illinois: Chicago or Springfield?). When deceptive questions were included in the test, participants' monitoring ability and performance suffered.

Receiving feedback may influence participants' ability to monitor their performance by participants using the feedback as a cue about their performance. When participants determine their confidence in performance they often use cues to judge how they are performing (e.g., Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002; Koriat, 1997). Koriat (1997) proposed the cue-utilization approach and found participants use different cues when predicting how well they have learned material. Receiving feedback regarding participants' performance should be a useful cue that improves how well they monitor their performance. Gonzalez-Vallejo and Bonham (2007) found that if participants were rewarded or punished after providing confidence judgments, their

monitoring became more accurate. Receiving feedback about Test 1 would be a strong cue about one's performance that participants may continue to use when assessing Test 2 performance. Receiving feedback may improve participants' test performance not only because they are receiving the correct answers, but also because they are receiving information to help them monitor and control their performance.

Current Experiments

The current studies were designed to examine the potential role monitoring has in influencing the effectiveness of feedback. More specifically, Experiment 1 was designed largely to refine our understanding of the role of confidence in the retention of correct responses. Several reviews of the effects of feedback reported that feedback has little impact on increasing the retention of correct responses (e.g., Kulhavy, 1977; Kulhavy & Stock, 1989), and many studies reported finding no consistent relationship between receiving feedback and the retention of correct responses (Kulhavy et al., 1990; McConnell & Hunt, 2007, Exp. 2; Pashler et al., 2005; Webb, Stock, & McCarthy, 1994). In contrast, Butler et al. (2008) examined the retention of correct responses, testing format, and confidence and demonstrated a relationship between them and the effectiveness of feedback. They explained that feedback enhances the retention of correct responses in the unique situation where participants are tested first with a multiple-choice test and later with a recall test. They suggest that feedback increases the retention of correct responses in this case because providing feedback after a recognition test makes participants more willing to recall those items later. In contrast, it is predicted that testing format is not the most important factor, but using procedures that encourage enough low

confident correct responses may be important in demonstrating an effect of feedback on the retention of correct responses. If participants answer enough questions correctly on Test 1 in low confidence, receiving feedback should enhance the retention of correct responses, even when using a multiple-choice format on Test 1 and Test 2. In Experiment 1, participants studied journal articles, answered multiple-choice questions about these articles, provided confidence ratings, received feedback on the last half of the questions, and later answered multiple-choice questions again. These materials were designed to be difficult enough that participants would have low confidence in even their correct responses.

The purpose of Experiment 2 was to determine whether participants detect and use false feedback. Participants' confidence in their original responses, and critically in the feedback itself, may influence the effectiveness of the feedback. Specifically, confidence in the feedback may moderate the effects of receiving it. Confidence in the feedback was manipulated by providing false feedback for some of the questions, and examining whether participants detected and continued to use the false feedback. In Experiment 2 participants read journal articles, answered multiple-choice test questions, provided confidence judgments, received feedback on the last half of the questions, and answered the test questions again. Importantly, 50% of the feedback participants received contained false information. Half of the false feedback confirmed a wrong response, similar to Hanaba and Zaragoza (2006), but the other half of the false feedback indicated a correct response was wrong. If participants are able to detect the false feedback, they may be less likely to use it to influence their performance. This will demonstrate whether

participants can detect false feedback after responding on Test 1, and if that influences whether they continue to use it.

Finally, Experiment 3 was designed to examine any differences between young children and college participants after receiving true feedback. Both 5- and 6-year-old and college participants heard a children's story, answered multiple-choice test questions, received feedback on performance, and completed another multiple-choice test. Preschool children may be less likely to benefit from receiving feedback compared to college students because of immature monitoring processes, information processing differences, or inclinations toward perseverative behavior. Alternatively, children may be limited to accessing the representation of the more recently received feedback. If they only have the representation of the feedback, they may be less likely to respond with original errors.

CHAPTER II

EXPERIMENT 1

The purpose of Experiment 1 was to examine if receiving feedback increases the retention of correct responses when participants answer Test 1 with enough low confident correct responses. It was also designed to test whether receiving true feedback improved performance in a laboratory setting that used materials similar to those found in typical classroom environments. Participants studied information, completed a test, received feedback on a subset of their responses, and completed the same test again. Previous research suggests participants' Test 2 performance should improve after receiving feedback (e.g., Kulhavy & Anderson, 1972). Participants also rated their confidence in their test responses, and research suggests participants' confidence ratings should be fairly accurate—they should be more confident in correct responses than errors (e.g., Koriat & Goldsmith, 1996). Participants' confidence in their performance is predicted to moderate the effects of feedback, with participants retaining low confident correct responses and correcting high confident errors (e.g., Butler et al., 2008; Butterfield & Metcalfe, 2001).

Method

Participants and Design

Ninety-six undergraduates from the University of North Carolina at Greensboro participated in the experiment to fulfill a course requirement. The participants came from Introduction to Psychology, 200–level psychology courses, and 300–level psychology courses.

Demographic information was not recorded.[1] The experiment was a within subjects design with two independent variable (Feedback: yes, no; and Test type: Test 1, Test 2).

Materials

The four *Current Directions in Psychological Science* articles used in Kang, McDermott, and Roediger (2007) were presented to participants in this experiment (see Appendix A). Participants read paper versions of the articles and completed a multiple-choice test after each article. Test 1 consisted of 4 subtests with each subtest corresponding to one article. Each subtest consisted of 12 multiple-choice questions with one correct response and 3 incorrect distractors; eight of the questions were used in Kang et al. and four more were added for this experiment (see Appendix B). The questions appeared on a computer with the 4 options in a predetermined order listed below the question. The participants read the articles in one of six different orders that were counterbalanced across participants. For the feedback, both questions answered correctly and incorrectly were displayed on the computer with the original question and the original options for 13 seconds, and then the correct response and the original response displayed below the question for 7 seconds (see Appendix C). Note that pilot testing suggested participants needed this amount of time to read the feedback. Test 2 consisted of the 48 Test 1 questions blocked by article but in a new random order.

Design and Procedure

Session 1. The first session lasted 2 hours. In groups of five or fewer, participants first completed a related experiment (see Experiment 3). Participants then had 12 minutes to read article 1. After reading the article they worked on a sudoku distractor puzzle for 3 minutes before completing an untimed test (Test 1) covering that article. The test questions were

presented on a computer screen one-at-a time, and participants provided their confidence on a 0 – 100 scale after each question. After completing the test questions, the procedures were repeated for articles 2, 3, and 4. After completing sub-test 1 for article 4, participants completed an unrelated task for 5 minutes. Participants then received feedback for the tests covering all the questions for articles 3 and 4 only. The feedback was always provided after completion of all the subtests. Feedback was always presented for articles 3 and 4 because previous findings demonstrated that the timing of the feedback influences its effectiveness (e.g., Hunt et al., 2009; Peeck & Tillema, 1978-1979; Phye, 1979; Phye & Andre, 1989). If feedback was provided after Articles 1 and 2 as well as 3 and 4, some of the feedback would be delayed while other feedback immediate. This may inadvertently have added another, unintended, variable to the studies (timing of feedback). Feedback was not provided after each test because previous findings demonstrate that when feedback is provided immediately on test questions when other test questions still remain to be answered, participants change their responding on Test 1 (Hunt et al., 2009). A concern with this design was that the feedback was never provided on the earlier questions, and participants may have performed more poorly on the no feedback test questions because they were less familiar with the task.

Session 2. Participants returned two days later and completed Test 2 consisting of the same 48 questions blocked by article but with questions presented in a new random order. As in session 1, the participants provided their confidence. Participants had an unlimited amount of time to complete the test in both sessions. A skeleton outline of the procedures for session 1 and session 2 includes: read article 1→ Test 1 article 1 → read article 2 → Test 1 article 2 → read

article 3→ Test 1 article 3 → read article 4 → Test 1 article 4 → brief delay → feedback over Test 1 for articles 3 and 4 → long delay → Test 2 over all articles.

Results

For all analyses, results are significant at the .05 level.

Overall Performance

Test 1 and Test 2 Proportion Correct. A 2 (Feedback: yes, no) X 2 (Test Type: Test 1, Test 2) repeated measures ANOVA on proportion correct revealed a main effect of test type, $F(1,95) = 8.775$, $p = .004$, $\eta_p^2 = .085$, a main effect of feedback, $F(1,95) = 30.090$, $p < .0001$, $\eta_p^2 = .242$ and a significant interaction between receiving feedback and test type, $F(1,95) = 79.030$, $p < .0001$, $\eta_p^2 = .454$ (see Figure 1). Follow up contrasts revealed no differences on Test 1 between questions that eventually received feedback and those that did not, $F < 1$, *ns*, but did reveal an effect of receiving feedback on Test 2 proportion correct, $F(1,95) = 9.326$, $p < .0001$, $d = .952$, such that participants answered more questions correctly after receiving feedback than after not receiving feedback.[2]

Conditionalized Test 2 Performance

Performance can also be analyzed by examining Test 2 performance conditionalized on Test 1 performance. Participants can answer questions correctly on both tests (retention of correct responses), answer questions incorrectly on both tests (retention of errors), answer questions incorrectly but change to a correct response (error correction), or they can answer questions correctly on Test 1 but change to an incorrect response on Test 2 (new errors).

Retention of responses. A repeated measures t-test revealed a significant effect of receiving feedback on the retention of correct responses, $t(95) = 2.617$, $p = .01$, $d = .267$, such that participants retained more Test 1 correct responses after receiving feedback than after not receiving feedback. Another repeated measures t-test also revealed a significant effect of receiving feedback on the proportion of errors retained, $t(95) = 10.870$, $p < .0001$, $d = 1.110$, specifically, participants retained fewer errors after receiving feedback than after not receiving feedback (see Table 1 for means and standard errors for responses that were retained).

Changing responses. A repeated measures t-test revealed a significant effect of receiving feedback on the proportion of errors corrected, $t(95) = 10.870$, $p < .0001$, $d = 1.457$, with participants correcting more Test 1 errors after receiving feedback than after not receiving feedback. Another repeated measures t-test revealed a significant effect of receiving feedback on the proportion of new errors, $t(95) = -2.617$, $p = .01$, $d = .280$, such that participants changed fewer Test 1 correct responses to errors after receiving feedback than after not receiving feedback (see Table 1 for means and standard errors for responses that were changed).

Confidence Ratings

The relationship between receiving feedback and participants' confidence in their test performance was also examined. Confidence ratings were an indicator of how well participants monitored their test performance. Due to a computer error, only Test 2 confidence ratings for 36 participants were recorded (i.e., data from 60 participants were not recorded).

Test 1 and Test 2 confidence ratings. A 2 (Feedback: yes, no) X 2 (Test type: Test 1, Test 2) repeated measures ANOVA on confidence ratings revealed no main effect of test type $F < 1$, *ns*, but did reveal a significant effect of feedback, $F(1,35) = 8.914$, $p = .005$, $\eta_p^2 = .203$, and a significant interaction between test type and receiving feedback, $F(1,35) = 10.300$, $p = .003$, $\eta_p^2 = .227$ (see Figure 2). Follow up contrasts revealed no preexisting differences for mean Test 1 confidence ratings, $F < 1$, *ns*, but did reveal that on Test 2 participants had higher confidence ratings in questions after receiving feedback than after not receiving feedback, $F(1,35) = 15.290$, $p = .0004$, $d = .650$.

Correct retention as a function of feedback and confidence. Previous research suggests that receiving feedback may have a different effect on the retention of correct responses and error correction for low and high confident responses. To that end, Test 1 confidence ratings from 0 – 29 were classified as low confidence, while confidence ratings from 80 – 100 were classified as high confidence[3]. A 2 (Feedback: yes, no) X 2 (Confidence rating: high, low) repeated measures ANOVA on correct retention revealed a main effect of receiving feedback, $F(1,41) = 6.089$, $p = .018$, $\eta_p^2 = .129$ a main effect of Test 1 confidence rating, $F(1,41) = 27.070$, $p < .0001$, $\eta_p^2 = .398$ and an interaction between receiving feedback and confidence rating, $F(1,41) = 12.367$, $p = .001$, $\eta_p^2 = .232$. Follow up contrasts revealed no effect of feedback on the retention of correct responses given in high confidence on Test 1, $F < 1$, *ns*, but did reveal an effect of feedback on Test 2 when participants answered a Test 1 response correctly in low confidence, $F(1,41) = 20.050$, $p < 0.0001$, $d = .660$ (see Figure 3). Specifically, participants were more likely to

retain low confident correct responses after receiving feedback than after receiving no feedback.

Error correction as a function of feedback and confidence. A 2 (Feedback: yes, no) X 2 (Confidence rating: high, low) repeated measures ANOVA on error correction revealed a significant main effect of receiving feedback, $F(1,31) = 6.483$, $p = .016$, $\eta_p^2 = .173$, such that participants retained fewer errors after receiving feedback than after not receiving feedback. There was no main effect of Test 1 confidence, $F(1,31) = 2.207$, $p = .147$, $\eta_p^2 = .066$, nor an interaction, $F < 1$, *ns* (see Figure 4). To test whether participants were more likely to correct errors made in high confidence than in low confidence after receiving feedback (the hypercorrection with feedback effect), an a priori repeated measures t-test examining feedback questions was conducted but revealed that participants did not correct more errors committed in high confidence ($M = .348$, $SE = .043$) than in low confidence after receiving feedback ($M = .490$, $SE = .074$), $t(56) = -1.685$, $p = .685$, $d = .450$.

Relationship between confidence and proportion correct. Gamma correlations are nonparametric statistics frequently used to measure metacognitive processes such as the relationship between confidence and test performance, and are commonly used when at least one of the variables is categorical. It is a relative measure of monitoring accuracy that ranges from +1.0 to –1.0, with positive gamma correlations reflecting that the participants were more likely to assign a higher confidence rating to correct responses than incorrect responses. The relationship between participants Test 1 confidence and Test 1 performance, Test 2 confidence and Test 2 performance, and Test 1 confidence

and Test 2 performance were examined as a function of receiving feedback. Single sample t-tests were used to determine whether these correlations were significantly different from zero, and revealed that participants monitoring was indeed significant in all cases, all $t(35)$'s > 5.815 and all p's < .0001. Recall that because of the computer error there were no confidence judgments for 60 participants on Test 2.

A 2 (Feedback: yes, no) X 2 (Test type: Test 1, Test 2) repeated measures ANOVA on the gamma correlations between test performance and confidence revealed no main effect of test type, $F(1,35) = 2.603$, $p = .116$, $\eta_p^2 = .069$, but did reveal a main effect of feedback, $F(1,35) = 5.229$, $p = .028$, $\eta_p^2 = .130$ and an interaction between feedback and test type, $F(1,35) = 9.590$, $p = .004$, $\eta_p^2 = .215$. Follow up contrasts revealed the gamma correlation on Test 1 did not differ betweens questions that eventually received feedback ($\gamma = .409$), and questions that did not eventually receive feedback ($\gamma = .387$), $F < 1$, *ns*. However, participants had higher gamma correlations on Test 2 for questions that received feedback ($\gamma = .526$), than questions that did not receive feedback ($\gamma = .387$), $F(1,35) = 13.903$, $p = .001$, $d = .621$.

Finally a repeated measures t-test revealed that the relationship between Test 1 confidence and Test 2 proportion correct was stronger for questions that did not receive feedback ($\gamma = .322$) than questions that did receive feedback ($\gamma = .221$), $t(94) = 2.427$, $p = .017$, $d = .212$.

Discussion

The results from Experiment 1 support the conclusion that feedback improves memory performance, and that the improvement was largely due to error correction.

Participants were also more likely to retain originally correct responses, less likely to turn correct responses to errors, and less likely to stick with originally incorrect responses after receiving feedback. Thus, consistent with previous findings, receiving feedback improved performance on Test 2 (e.g., Kulhavy & Anderson, 1972; McConnell & Hunt, 2007; Pashler et al., 2005).

Not only did feedback influence test performance, but it also influenced participants' confidence in their performance. Participants were more confident after receiving feedback than after not receiving feedback, suggesting that participants recognized the potential benefit of receiving feedback. Importantly, confidence ratings moderated the effects of feedback, as seen when participants retained a greater proportion of Test 1 low confident correct responses after receiving feedback, but they did not retain a greater proportion of Test 1 high confident correct responses after receiving feedback. These findings support Butler et al.'s (2008) conclusions that feedback can enhance the retention of correct responses, and that this benefit is apparent even when using different testing procedures than in Butler et al.'s study. These findings suggest feedback may enhance the retention of correct responses in a broader range of contexts than Butler et al. first suggested. One interpretation of this finding is that many of the low confident correct responses can be thought of as correct guesses, and receiving confirmatory feedback may confirm these guesses increasing the likelihood of answering those questions correctly again (Butler et al., 2008; Hunt et al., 2009, Exp. 1).

Surprisingly, participants did not correct more errors committed in high confidence after receiving feedback (e.g., Butterfield & Metcalfe, 2001; Kulhavy et al.,

1979). Butterfield and Metcalfe (2001) explained the hypercorrection with feedback effect by assuming participants were surprised the high confident errors were incorrect, so they paid more attention to that feedback. Similarly, Kulhavy et al. (1979) found that participants spent more time studying feedback for errors made in high confidence than low confidence indicating that when participants allocate how much time they process the feedback they may allocate more time to feedback they find surprising. The hypercorrection with feedback effect may not have occurred in the current study because participants did not control how much time they spent studying the feedback—all participants read the feedback for 20 seconds, and they may have been unable to spend extra time on feedback that surprised them. This suggests the hypercorrection with feedback effect may occur only when using methods that allow participants enough time to study the feedback that they can allocate more time to some feedback than other feedback. Another possible reason for the failure to replicate the hypercorrection with feedback effect is the studies that typically report the effect have short retention interval (e.g., Butterfield & Metcalfe, 2001, 2006), while the studies that have failed to find an effect had longer retention intervals (e.g., Butler et al., 2008; Pashler et al., 2005).

Receiving feedback also influenced the relative accuracy of the confidence ratings. Consistent with previous findings (e.g., Koriat & Goldsmith, 1996) there were positive gamma correlations between participants' confidence and performance; furthermore, participants better monitored their performance on Test 2 after receiving feedback than after not receiving feedback. Finally, participants' Test 1 confidence ratings were related to their Test 2 proportion correct, but this relationship was stronger

when they did not receive feedback. This is unsurprising, because when participants receive outside information (i.e., the feedback) they have other information to influence how they respond on Test 2. These results demonstrate that participants may use feedback to assess their performance, as when they received feedback their relative accuracy improved. This finding confirms other research demonstrating participants use cues when judging their performance (e.g., Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002; Koriat, 1997).

Participants used the feedback effectively and also appeared to recognize the benefit of receiving feedback by indicating higher confidence. Although participants use the feedback when the information should improve performance, it is unclear if participants could ignore the feedback when it would impair performance. If participants use the feedback because it improves performance, then they should ignore false feedback. Experiment 2 investigates whether participants continue to use feedback that contains false information.

CHAPTER III

EXPERIMENT 2

Experiment 2 was designed to assess whether participants detect false feedback, and if they do, whether participants can ignore it. This is important because although participants' performance in Experiment 1 indicates that confidence in their original responses influences how they use the feedback, it is unclear whether participants' confidence in the feedback influences how they use it. If participants are less confident in the feedback, then they may be more likely to respond with their original responses, which would indicate that participants' metacognitive assessment of the situation influences how they use feedback. Research investigating the hindsight bias might suggest that participants are unable to ignore feedback. The findings from several hindsight bias studies reveal participants use feedback to influence their memory performance, even when told to ignore the information (Fischhoff, 1977; Fischhoff & Beyth, 1975; Hell et al., 1988; Pohl, 1998; Pohl & Hell, 1996). In an impressive demonstration of the hindsight bias, hindsight bias researchers answered trivia questions and received the answers. When asked for their original responses, these expert participants demonstrated the hindsight bias, and it persisted even after they were shown they were demonstrating the bias (Pohl & Hell, 1996).

Hasher et al. (1981), however, demonstrated participants are able to ignore information when explicitly warned the feedback contains false information. In the current study, only 50% of the feedback was correct. Importantly, unlike in previous studies, participants received no

warning that any of the feedback they received may be incorrect, and it is unclear whether participants can detect and ignore false feedback without receiving a warning. The findings from several studies indicate that participants may notice the false feedback if it contradicts their existing knowledge, because participants have direct access to whether their answers were correct (Hart 1965, 1966, 1967; Nelson & Dunlosky, 1994). For example, Hart claimed participants have an internal monitor that can detect when the correct information is stored in memory and when it is not. Hart was interested in how participants know if they have access to correct stored information, even if they cannot currently access it. He used a recall-judgment-recognition paradigm where participants were asked to recall answers to general knowledge questions. If they could not recall the answer, they assessed how likely they would be able to recognize the answer (i.e., feeling of knowing judgments). Hart demonstrated that participants feeling of knowing judgments were highly correlated with participants' ability to later recognize the answer. He also suggested that participants' feeling of knowing judgments were higher when they accessed the correct answer than an unrelated incorrect answer. If participants have direct access, then the false feedback would contradict their stored knowledge and they may reject the feedback. Other research suggests that participants may not notice the false feedback because they do not have direct access to whether their answers were correct (e.g., Koriat, 1993; Metcalfe, Schwartz, & Joaquim 1993). The results of Koriat's (1993) research, for example, reveals that participants judge how likely their answers were correct from inferences such as how quickly they retrieved their answer. When participants retrieved their answer quickly, they judged it was correct; on the other hand, when it took longer to retrieve their answer they judged it was incorrect.

In the current study, participants either received true feedback, false feedback, or no feedback for each question after completing Test 1. Participants who noticed some of the feedback was incorrect may have ignored the false feedback. If this were true, performance on the true feedback questions would be better than on the false feedback and no feedback questions, but performance would not differ between the false feedback and no feedback questions (true feedback > no feedback = false feedback). On the other hand, participants might have noticed some of the feedback was false and thus ignored the false and the true feedback. If this were true, performance would look similar across true feedback, false feedback, and no feedback questions (true feedback = no feedback = false feedback). Another possibility is participants may have noticed some of the feedback was false but were unable to ignore the false feedback. In this scenario, the performance would be best in the true feedback condition, second best when no feedback was received, and worst when false feedback was received (true feedback > no feedback > false feedback). The final possibility is that participants may not have noticed the feedback was false. If this were true, then the performance would follow the same pattern as when participants noticed but were unable to ignore the false feedback (true feedback > no feedback > false feedback).

Methods

Participants and Design

Ninety undergraduates from the University of North Carolina at Greensboro participated in this repeated measures design to fulfill a course requirement. The participants come from the same population as Experiment 1. The completely within design included one

independent variable with three levels (Feedback type: true, false, no feedback), and one independent variable with two levels (Test type: Test 1, Test 2).

Materials and Procedure

The same materials (journal articles and test questions) and the same procedure order from Experiment 1 were used in Experiment 2. There were four differences in the procedure:

1. Fifty percent of the feedback was false. Half of the false feedback indicated the incorrect answer was correct, whereas the other half of the false feedback indicated the correct answer was incorrect. All of the questions had a predetermined incorrect answer designated for the false feedback after a participant answered correctly. The computer assigned the first three questions answered correctly and the first three questions answered incorrectly as the false feedback. Although false feedback was always provided on the first three questions answered correctly and incorrectly, the order of the Test 1 questions were random, ensuring the false feedback was not provided on the same questions for all participants. Also, the feedback was randomly displayed throughout the feedback session, such that even though the false feedback was assigned first, it was not displayed in the same order the Test 1 questions originally appeared, but was randomly displayed throughout the feedback session.

2. To ensure half of the false feedback incorrectly confirmed a wrong answer and half of the false feedback incorrectly indicated a that correct answer was wrong, participants had to answer a certain amount of questions incorrectly on Test 1 for articles 3 and 4. Only participants who answered at least three questions correctly and incorrectly on Test 1 for articles 3 and 4 received the appropriate proportion of both types of false feedback and were used in the analyses (n = 13 participants were removed).

3. Due to computer programming, the top half of the feedback screen displayed the original question and multiple choice options, and the bottom half of the screen displayed *only* the letter of the original response and the letter of the correct response (see Appendix D). Because *only the letter* (and not the entire answer) was displayed, feedback was shown for two more seconds to allow participants time to glance up to the original options.

4. Participants did not first participate in another experiment.

Results

Overall Performance

Test 1 proportion correct. Because 50% of the feedback was false and divided equally into the 2 types of false feedback, all participants used in these analyses answered 50% of Test 1 false feedback questions correctly. Therefore, to ensure no baseline differences between true feedback and no feedback questions, Test 1 proportion correct was analyzed comparing true feedback and false feedback only. A paired samples t-test revealed no preexisting differences between questions that eventually received true feedback (M = .5510, SEM = .0270) and no feedback (M = .5550, SEM = .0170), $t(66) = -.169$, $p = .866$, $d = .041$.

Test 2 proportion correct. A repeated measures ANOVA revealed a significant effect of feedback type on Test 2 proportion correct, $F(2,132) = 93.301$, $p < .0001$, $\eta_p^2 = .586$. A priori contrasts revealed when participants received true feedback ($M = .649$, $SE = .024$) they answered more Test 2 questions correctly than when they received no feedback ($M = .499$, $SE = .019$), $F(1,66) = 50.183$, $p < .0001$, $d = 1.74$, which in turn yielded more correct responses than false feedback ($M = .289$, $SE = .019$), $F(1,66) = 62.929$, $p < .0001$, $d = 1.95$ (true feedback > no feedback > false feedback).[4]

Because Test 1 and Test 2 proportion correct was analyzed separately, participants' improvement between Test 1 and Test 2 was examined to determine how receiving feedback influenced performance across Tests 1 and 2. Test improvement was measured by calculating Test 2 proportion correct – Test 1 proportion correct. Single sample t-tests revealed participants' performance improved when they received true feedback, $t(66) = 3.921$, $p < .0001$, but when they received no feedback and false feedback their performance suffered, $t(66) = -4.274$, $p < .0001$, and $t(66) = -11.709$, $p < 0.0001$, respectively. A repeated measures ANOVA revealed a significant effect of feedback type on the percent improvement score (Test 2 proportion correct – Test 1 proportion correct), $F(2,132) = 59.416$, $p < .0001$, $\eta_p^2 = .474$. A priori contrasts revealed when participants received true feedback ($M = 9.8\%$, SE = 2.5) their performance improved more than when they received no feedback ($M = -5.5\%$, $SE = 1.5$), $F(1,66) = 27.042$, $p < .0001$, $d = 1.28$ which in turn yielded less decline than when they received false ($M = -21.9\%$, $SE = 1.9$), $F(1,66) = 41.917$, $p < .0001$, $d = 1.59$ (true feedback > no feedback > false feedback).

Retention of responses. A repeated measures ANOVA revealed a significant effect of feedback type on the retention of correct responses, $F(2,132) = 69.805$, $p < .0001$, $\eta_p^2 = .514$ (see Table 2). A priori contrasts revealed when participants received true feedback they retained a greater proportion of Test 1 correct responses than when they received no feedback, $F(1,66) = 6.066$, $p = .016$, $d = .605$, which in turn yielded more retention than when they received false feedback, $F(1,66) = 85.439$, $p < .0001$, $d = 2.27$, (true feedback > no feedback > false feedback).

A repeated measures ANOVA revealed a significant effect of feedback type on the retention of Test 1 errors, $F(2,132) = 52.962$, $p < .0001$, $\eta_p^2 = .449$ (see Table 2). A priori contrasts revealed when participants received true feedback they retained a smaller proportion of Test 1 errors than when they received no feedback, $F(1,66) = 52.760$, $p < .0001$, $d = 1.78$, which in turn yielded less retention than when they received false feedback, $F(1,66) = 7.445$, $p = .008$, $d = .672$ (true feedback < no feedback < false feedback).

Changing Responses. A repeated measures ANOVA revealed a significant effect of feedback type on error correction, $F(2,132)$ 52.962, $p < .0001$, $\eta_p^2 = .449$ (see Table 2). A priori contrasts revealed when participants received true feedback they corrected a greater proportion of Test 1 errors than when they received no feedback, $F(1,66) = 52.760$, $p < .0001$, $d = 1.78$, which in turn yielded more error correction than when they received false feedback $F(1,66) = 7.445$, $p = .008$, $d = .672$ (true feedback > no feedback > false feedback).

A repeated measures ANOVA revealed a significant effect of feedback type on new errors, $F(2,132) = 69.805$, $p < .0001$, $\eta_p^2 = .509$. A priori contrasts revealed when participants received true feedback they changed a smaller proportion of Test 1 correct responses to Test 2 errors than when they received no feedback, $F(1,66) = 6.066$, $p = .016$, $d = .605$, which in turned yielded fewer new errors than when they received false feedback, $F(1,66) = 83.727$, $p < .0001$, $d = 2.27$, see Table 2 (true feedback < no feedback < false feedback).

Proportion of Responses that Match the Feedback

To assess how likely participants were to use the true and false feedback, performance was analyzed by examining proportion of Test 2 responses that matched the true and false feedback. A 2 (Feedback type: true, false) X 2 (Test 1 correct: correct, incorrect) repeated measures ANOVA revealed no main effect of whether the Test 1 response was answered correctly, $F < 1$, *ns*, but did reveal a significant main effect of feedback type, $F(1,65) = 14.385$, $p < .0001$, , $\eta_p^2 = .181$, and an interaction between feedback type and Test 1 correctness, $F(1,65) = 73.742$, $p < .0001$, , $\eta_p^2 = .532$ (see Figure 5). Follow up contrasts revealed that when participants answered questions correctly on Test 1 they were more likely to follow the true feedback than the false feedback, $F(1,65) = 69.483$, $p < .0001$, $d = 1.02$, but when participants answered the question incorrectly on Test 1 they were more likely to follow the false feedback than the true feedback, $F(1,65) = 7.135$, $p = .01$, $d = .329$. These results suggest that participants tended to retain their original responses, and follow the feedback that confirms their original response.

Confidence Ratings

Another indirect method to examine whether participants noticed some of the feedback was false was to compare Test 2 confidence in answers that received true feedback, false feedback, and no feedback. Specifically, if participants noticed the false feedback, then they may be less confident in these responses.

Mean confidence rating as a function of feedback type and test type. A 3 (Feedback type: true, false, no feedback) X 2 (Test type: Test 1, Test 2) repeated

measures ANOVA revealed no main effect of feedback type on mean confidence rating, $F < 1$, *ns*, no main effect of test type, $F(1,27) = 1.262$, $p = .335$, and no significant interaction, $F < 1$, *ns* (see Figure 6).

The retention of correct responses as a function of feedback type and confidence. The confidence participants have in their original responses influence the effectiveness of the feedback (e.g., receiving feedback effectively enhanced the retention of correct responses for Test 1 low confident responses only). Examining whether confidence in responses still influences the effectiveness of feedback when some of the feedback is incorrect may further explain the relationship between confidence and feedback. When participants receive false feedback indicating a highly confident correct response was false, the participant may be less likely to trust the content of the feedback. A repeated measures ANOVA revealed a significant effect of feedback type on the retention of Test 1 high confident (80 –100) correct responses, $F(2,78) = 88.705$, $p < .0001$, $\eta_p^2 = .695$. Follow up contrasts revealed participants were no more likely to retain high confident correct responses after receiving true feedback than after receiving no feedback, $F < 1$, *ns*. In contrast, participants were more likely to retain high confident correct responses after receiving true feedback than after receiving false feedback, $F(1,39) = 124.634$, $p < .0001$, $d = 3.57$, and no feedback than false feedback, $F(1,39) = 138.255$, $p < .0001$, $d = 3.57$.

A repeated measures ANOVA also revealed an effect of feedback type on the retention of correct responses given in low confidence (0 – 29), $F(2,36) = 11.617$, $p < .0001$, $\eta_p^2 = .392$. Follow up contrasts revealed participants were more likely to retain low

confident correct responses after receiving true feedback than after receiving no feedback, F(1,18) = 10.920, p = .004, d = 1.42, and after receiving true feedback than false feedback, F(1,18) = 18.500, p < .0001, d = 2.02. Finally, participants were not more likely to retain low confident correct responses on questions that did not receive feedback than questions that received false feedback, F(1,18) = 2.751, p = .114, d = .78 (see Figure 7).

Error correction as a function of feedback type and confidence. A repeated measures ANOVA revealed a significant effect of feedback type on the correction of high confident Test 1 errors, $F(2,50) = 10.770$, p = .004, η_p^2 = .301. Follow up contrasts revealed participants corrected more high confident errors after receiving true feedback than no feedback, $F(1,25) = 13.832$, p = .001, d = 1.48, and false feedback $F(1,25)$ = 12.100, p = .002, d = 1.33. There was no difference between questions that did not receive feedback and received false feedback, F < 1, *ns*. An a priori contrast revealed no significant differences in error correction of low confident Test 1 errors as a function of feedback type, F < 1, *ns* (see Figure 8). A repeated measures t-test also revealed no hypercorrection with feedback effect. Participants were as likely to correct Test 1 errors made in high confidence (M = .390, SE = .102) as errors made in low confidence (M = .270, SE = .090) after receiving true feedback, $t(16) = .846$, p = .410, d = .423.

Relationship between confidence and proportion correct. Single sample t-tests comparing gamma correlations for the relationship between participants Test 1 confidence and Test 1 performance, Test 2 confidence and Test 2 performance, and Test 1 confidence and Test 2 performance as a function of feedback type to zero revealed

significant monitoring, all $t(26)$'s > 4.514, all p's < .067, except for the relationship between Test 2 confidence ratings and Test 2 performance after receiving false feedback, t(26) = -.002, p = .983. Low and High Confidence correct retention was analyzed in separate ANOVAs because only 8 participants had performance that allowed them to be analyzed together.

A 3 (Feedback type: true, false, no feedback) X 2 (Test type: Test 1, Test 2) repeated measures ANOVA on the gamma correlations between test performance and confidence revealed no main effect of feedback, $F(2,44) = 2.342$, p = .108, η_p^2 = .096, and no main effect of test type, F <1, ns, but did reveal an interaction between feedback and test type, $F(2,44) = 8.839$, p = .001, η_p^2 = .287. Follow up contrasts revealed that participants' gamma correlation did not differ on questions as a function of feedback type on Test 1, $F < 1$, *ns*, η_p^2 = .010, true feedback (γ = .403) = no feedback (γ = .385) = false feedback (γ = .331). Participants' gamma correlations did differ on questions as a function of feedback type on Test 2, $F(2,44) = 8.848$, p = .001, η_p^2 = .473. Specifically, a priori contrasts revealed that participants had higher gamma correlations on Test 2 for questions that received true feedback (γ = .558) than questions that received no feedback (γ = .257), $F(1,22) = 6.724$, p = .017, d = 1.10 , and they had higher gamma correlations for questions that received true feedback than false feedback (γ = -.002), $F(1,22)$ = 19.506, $p < .0001$, d = 8.313.Contrasts also revealed participants did not have higher gamma correlations after receiving no feedback than false feedback, $F(1,22) = 1.957$, p = .176, d = .834. Low and High Confidence error correction was analyzed in separate

ANOVAs because only 7 participants had performance that allowed them to be analyzed together.

Single sample t-tests comparing gamma correlations to zero did not reveal a significant relationship between Test 1 confidence and Test 2 performance after receiving false feedback, $t(65) = 1.750$, $p = .085$, but did after receiving true feedback and no receiving feedback, all $t(65)$*'s* > 3.780, all p*'s* $< .0001$.[5]

A repeated measures ANOVA (Feedback type: true, false, no feedback) revealed an effect of feedback type on the relationship between Test 1 confidence and Test 2 proportion correct, $F(2,126) = 3.538$, $p = .032$, $\eta_p^2 = .110$. A priori contrasts revealed a marginal effect of feedback type with a weaker relationship between Test 1 confidence ratings and Test 2 proportion correct after receiving true feedback ($\gamma = .260$) than no feedback ($\gamma = .312$), $F(1,63) = 3.466$, $p = .067$, $d = .468$, but the relationship was similar after receiving true feedback and false feedback ($\gamma = .113$), $F < 1$, *ns*. Finally, the relationship between Test 1 confidence and Test 2 proportion correct was stronger after receiving no feedback than false feedback, $F(1,63) = 5.65$, $p = .008$, $d = .598$.

Awareness

After completing the experiment, participants reported whether they noticed anything about the feedback, and 15 participants indicated they noticed the feedback did not match the content in the articles. Participants' test performance and confidence ratings as a function of feedback type were analyzed after removing participants who reported being aware of the false feedback. Not surprisingly, analyses revealed the same patterns when these participants were removed as when they were included. Participants'

test performance and confidence ratings were also analyzed treating awareness as an independent variable. All results were analyzed as a function of awareness, and the following patterns discussed in the results were significant, but a full analysis of all the independent variables was not possible due to the small number of aware participants. Therefore, the patterns and not the statistics are reported. Participants who reported being aware answered more questions correctly on Test 1 and maintained the enhanced performance on Test 2 (see Figure 9). These participants were also more likely to retain correct responses and less likely to change correct responses to errors. They were, however, just as likely to retain errors and correct errors as were participants who were unaware (see Table 3). Participants who reported being aware some of the feedback was false were more likely to follow the true feedback when they answered correctly than unaware participants, but they were just as likely to follow the false feedback as unaware participants (see Figure 10). When participants reported being aware answered incorrectly, they were just as likely as the unaware participants to follow the true and false feedback. Finally, aware participants were more confident on Test 1 than were unaware participants. Only 2 participants who reported being aware of the false feedback had Test 2 confidence ratings, so interpreting how awareness influenced Test 2 confidence is impossible.

Discussion

An examination of performance based on the type of feedback indicated participants used both the true feedback and the false feedback. Specifically, participants' performance improved when they received true feedback but suffered when they received

false feedback. Examining performance alone does not indicate whether participants detected if the feedback was false. Participants may have detected the false feedback but have been unable to ignore the feedback when answering Test 2 questions. One method of assessing whether participants detected the false feedback was to examine the participants' confidence in their performance on the second test. Accordingly, participants did not detect the false feedback because (a) they were as confident in the false feedback responses as the other types, and (b) Test 2 confidence ratings were similar to Test 1. This suggests that providing false feedback does not influence participants' confidence in their performance, at least as long as they are unaware the feedback provided false information. This is unsurprising because if participants were unaware the feedback is false, then there would be no reason to expect their confidence to lower.

Receiving false feedback also influenced participants' monitoring ability. Participants demonstrated relative accuracy in their monitoring of their test performance, such that on both Test 1 and Test 2, participants were more confident in their correct responses. However, participants who received false feedback did not demonstrate a higher relative accuracy in their monitoring performance on Test 2. This demonstrates that receiving misinformation about one's performance disrupts one's ability to monitor test performance, and it illustrates the difficulties individuals can experience when monitoring their performance. This finding is consistent with Koriat and Goldsmith's (1996) conclusions that participants can monitor effectively in certain situations, but demonstrate monitoring difficulties in others.

Interestingly, the results also revealed participants were more likely to use the feedback that indicated they should keep their original responses, regardless of the correctness of the feedback. Examining the participants' test performance and confidence suggests participants used the false feedback and may not have noticed the feedback was false. These findings are in line with the findings of Hanaba and Zaragoza's (2006) investigation of false feedback and eyewitness testimony. In their study, participants were told their wrong answers in an eyewitness testimony paradigm were actually correct and were thus more likely to continue to believe that their wrong answers were correct. No other research has provided false feedback when participants were originally correct. If participants somehow had inside knowledge of when their responses were correct, then they would have ignored the false feedback that indicated their correct responses were wrong (e.g., Hart, 1965). But, because participants used the false feedback and changed correct responses to errors, it indicates participants may not have had special insight into their responses, but used outside information and inferences to determine when they were correct (e.g., Koriat, 1993, 2008).

Few participants indicated they noticed that some of the feedback was false, and these participants started out with enhanced performance and maintained it on both tests. These findings may indicate that when participants were aware of the false feedback they were able to ignore the false feedback and stick with their original answers. Alternatively, participants who were aware of the false feedback may be more engaged, and thus more likely to use the true feedback to better their performance. The findings indicated that aware participants had better performance because they were engaged and more likely to

follow the true feedback. Still, even though they noticed some of the feedback was false, they were no less likely to follow the false feedback than were the unaware participants. All participants' performance would have suffered less if they ignored all of the feedback, but only the aware participants had reason to try and ignore the feedback. In fact, several aware participants responded that they tried to ignore the provided feedback and respond with their original answers. As evidenced by their performance, aware participants did not ignore the feedback, which suggests participants may have difficultly discriminating their original responses from the provided feedback.

CHAPTER IV

EXPERIMENT 3

The goal of Experiment 3 was to examine the effects of receiving feedback in preschool children and college students. Experiment 3 was a first step in examining whether preschool children benefit from receiving feedback on their test performance. Five- and 6-year-old children heard age appropriate stories, answered questions, and rated their confidence in their responses. After answering the questions, the children received feedback (all true) on half of their answers and then after a delay answered the same questions again.

Sixty-two of the college participants from Experiment 1 also participated in this study. There are several reasons to predict preschool participants will show less of a benefit from the feedback. As mentioned previously, children and adults do not consistently benefit from the same memory strategies. Specifically, young children may use a strategy that benefits adults memory performance, but does not improve the children's performance because the strategy uses cognitive resources needed for other components of the task (Bjorkland & Harnishfeger, 1987; Miller et al., 1994). Not only do children fail to benefit from certain strategies, but their memory organization also differs from college students. Preschool children do not process the relationship among items as easily as older children and college students. Although young children can categorize information, often the thematic information needs to be very salient for it to help the children (Brainerd & Reyna, 2004; Denney, 1974; Kulig & Tighe, 1976; Schneider & Bjorklund, 1998). In previous studies, preschool children have been found to

classify information on perceptual similarity and not thematic similarity (Denney, 1974), they demonstrate less category clustering during free recall (Schneider & Bjorklund, 1998), and as children memorize categorized lists often they fail to detect the theme (Kulig & Tighe, 1976). Because children's memories are organized less thematically than adults, it may be more difficult to integrate the feedback into their memories; thus the feedback may be less likely to come to mind and would prove less helpful. On the other hand, from a source monitoring perspective, children may benefit similar as adults after receiving feedback because of good external source monitoring skills (e.g., Drummey & Newcombe, 2002). But, if the feedback and the original response are very similar, children may benefit less than college participants after receiving feedback because of enhanced difficulties discriminating similar sources (Lindsay et al., 1991).

Regardless, there is reason to predict that young children may benefit from receiving feedback. First, children have the ability to discriminate between two external sources (e.g., Drummey & Newcombe, 2002; Lindsay, Johnson, & Kwon, 1991; Powell & Thomson, 1997), indicating that children may easily discriminate the Test 1 original response from the feedback. Second, children demonstrate the hindsight bias (e.g., Bernstein, Atance, Metlzoff, & Loftus, 2007; Birch & Bernstein, 2007), which suggests their memory performance is influenced by receiving feedback. Bernstein et al. (2007) had 3- to 5 -year-old preschool children view degraded objects as the images gradually became focused, and they named the object as soon as it became focused enough to identify it. After finishing this task the children viewed the same degraded objects and were asked to identify when a friend would be able to identify the

objects. The children demonstrated an ability to benefit from feedback by identifying the objects for a friend more quickly than the children originally identified the objects. These findings demonstrate that receiving a form of feedback can influence preschool children's performance. Also, as mentioned previously, preschool children have the ability to discriminate between two external sources of information (Foley et al., 1983). They may remember both the original response and the feedback when answering Test 2, and successfully discriminate between responses. Also, young children may benefit from receiving feedback because they may only remember the feedback. As Templeton and Wilcox (2000) suggested, young children may only remember the more recent source of information because they can only hold one representation.

Although the college participants completed the experiment using the same materials as the preschool children, the delay between receiving the feedback and Test 2 differed for the preschool and college participants. The preschool participants completed Test 2 ten minutes after receiving the feedback, while the college participants completed Test 2 forty-eight hours after receiving the feedback. This procedural difference was necessitated because of previous research suggesting preschool children have a faster forgetting rate than older children (e.g., Brainerd & Reyna, 1990,1995; Howe, 1991). For example, Howe examined forgetting rates at two different delay intervals (2 days and 9 days) at two different ages (kindergarten and 2nd grade), and in his first experiment he found that kindergarteners recalled fewer items at a 2-day delay than 2nd graders did at a 9-day delay. Because the difference in delay, any differences found between college and preschool participants need to be interpreted with caution.

Methods

Participants and design. Sixty-two college students from Experiment 1 completed the experiment. Also fifty 5- and 6- year-old preschool children recruited from local daycares, preschools, little league sporting events, and a mailing list of local children participated in this experiment after receiving permission from their parents. Results were not obtained for 6 preschool participants because they did not complete the experiment or because of experimenter error. The final preschool sample consisted of 24 females (mean age = 6.063 and SD = .629) and 20 males (mean age = 6.100 and SD = .590). The participants completed this mixed subjects design with two within subjects independent variables (Feedback: yes, no; and Test type: Test 1, Test 2) and one between subjects independent variable (Age: college, preschool).

Preschool Children Methods

Materials. The materials included two stories (*Danny and the Dinosaur* by Sid Hoff and *Nobody Notices Minerva* by Wednesday Kirwan), and a multiple-choice test consisting of 12 questions for each story. The multiple-choice questions and options were illustrated, and the questions and options were presented in a fixed order for all participants. The multiple-choice illustrations for each question were on a piece of paper and depicted the main noun or action in the answer. Because not all of the children were able to read, the question and options were not typed out below the illustrations, see Appendix E. The questions were piloted to have children answering around 60% of the questions correctly.

A confidence rating scale was also used. The scale was a 3-point Likert scale intended for children consisting of three smiley faces with varying degrees of a smile (smile, neutral,

frown) that represented levels of confidence (Roebers, 2002), see Appendix F. The smile face indicated "sure or remember it", the neutral face indicated "kind of sure", and the frown face indicated "just guessing". The feedback after a correct response consisted of the correct response marked by a star, and after an incorrect response the feedback consisted of the correct response marked by a star and the wrong answer marked with an X. Test 2 consisted of the same questions blocked by story and in an identical order as Test 1.

Procedure. The procedure was designed to parallel Experiment 1. Children participated one at a time and were first trained to use the confidence scale with the smiley faces (Roebers, 2002). During the training session, the children learned what each smiley represented and then answered questions unrelated to the children stories. All children were asked the same questions during training, and if they had difficulty answering the questions the experimenter provided other examples for extended training. Following training, participants heard Story A (either *Danny and the Dinosaur* or *Nobody Notices Minerva)*, and then played various games during a 5 minute delay. They then answered 12 multiple-choice questions and provided their confidence after each question. Five minutes after completing the test they heard Story B, and after another 5-minute delay they answered and gave their confidence on 12 multiple-choice questions covering Story B. The children played games for 5 minutes and then received feedback on the test for Story B only. Feedback was provided by the experimenter marking the responses and then pointing to the child's answer and saying "This is your answer." If the answer was correct, the experimenter said, "Let's see the answer. Good job! You are right. You said x and it was x. What is the right answer?", but if the answer was incorrect, the experimenter said, "Let's see the answer. No, that is not right. You said X, but see this is the

right answer. What is the right answer?" Ten minutes after receiving feedback children completed an identical Test 2 and provided their confidence ratings.

College Participants Methods

Materials. The same stories, test questions, order of the test questions, and confidence scale used with the preschool participants were used with the college participants. There was one difference in the methods, and it was that the question and written descriptions of the options were typed on each test page (see Appendix G). The college participants received feedback only on half of the test questions, and the feedback questions were determined by one of six different orders that were counterbalanced across participants.

Procedure. In groups of five or fewer, participants heard only one of the two stories read by the experimenter, and after hearing it college students participated in Experiment 1 for about an hour and 15 minutes. Participants then read and answered the 12 test questions, and after a 5-minute delay they received feedback on half of their questions and completed an identical Test 2 forty-eight hours after completing Test 1.

Results

There were no effects of the sex of preschool participants for any of the analyses, nor did sex interact with any predictor variable. Sex will not be discussed further.

Test 1 and Test 2 Proportion Correct

A 2 (Feedback: yes, no) X 2 (Test type: Test 1, Test 2) X 2 (Age: college, preschool) repeated measures ANOVA on proportion correct revealed a main effect of receiving feedback, $F(1,104) = 51.243$, $p < .0001$, $\eta_p^2 = .330$, a main effect of test type, $F(1, 104) = 192.238$, $p < .0001$, $\eta_p^2 = .649$, and a main effect of age, $F(1,104) = 5.285$, p

= .024, η_p^2 = .024. There was also a significant interaction between feedback and age, $F(1,104) = 8.340$, $p = .005$, $\eta_p^2 = .074$ and an interaction feedback and test type, $F(1,104) = 80.631$, $p < .0001$, $\eta_p^2 = .437$. There was no significant interaction between test type and age, $F(1,104) = 1.717$, $p = .193$, $\eta_p^2 = .016$. These interactions were qualified by a three-way interaction between feedback, test type, and age, $F(1,104) = 4.050$, $p = .047$, $\eta_p^2 = .037$ (see Figure 11). To dissect the three-way interaction, performance was divided by Test 1 and Test 2, and a simple effects test examining Test 1 performance revealed a main effect of age, such that college participants answered a greater proportion of questions correctly than did preschool participants, $F(1,104) = 5.203$, $p = .025$, $\eta_p^2 = .048$, but no main effect of feedback, $F < 1$, *ns*, and no interaction, $F(1,104) = 1.298$, $p = .257$, $\eta_p^2 = .012$. A simple effects test examining Test 2 performance revealed a main effect of feedback, such that a greater proportion of feedback questions were answered correctly than no feedback questions, $F(1,104) = 120.833$, $p < .0001$, $\eta_p^2 = .538$, and an almost significant effect of age, $F(1,104) = 3.804$, $p = .054$, $\eta_p^2 = .035$, with college participants answering a greater proportion of questions correctly. This was qualified by a significant interaction between feedback and age, $F(1,104) = 12.167$, $p = .001$, $\eta_p^2 = .105$. Follow up contrasts revealed there were no differences on feedback questions as a function of age, $F(1,104) = 1.348$, $p = .181$, $d = .263$, but on questions that did not receive feedback, college participants answered more questions correctly than did preschool participants, $F(1,104) = 3.239$, $p = .002$, $d = .635$.

Retention of Responses

A 2 (Feedback: yes, no) X 2 (Age: college, preschool) repeated measures ANOVA on the proportion of correct retention revealed a main effect of feedback, $F(1,104) = 10.391$, $p = .002$, $\eta_p^2 = .092$, an almost significant main effect of age, $F(1,104) = 3.249$, $p = .074$, $\eta_p^2 = .074$, and a significant interaction between feedback and age, $F(1,104) = 4.78$, $p = .034$, $\eta_p^2 = .044$. Follow up contrasts revealed no effect of feedback on the proportion of correct retention for college students, $F < 1$, *ns*; whereas, there was an effect of feedback for preschool participants, such that they retained more correct responses after receiving feedback than after not receiving feedback, $F(1,104) = 8.49$, $p = .006$, $d = 1.665$ (see Table 4 for means and standard errors of retained responses).

A 2 (Feedback: yes, no) X 2 (Age: college, preschool) repeated measures ANOVA on the proportion of error retention revealed no main effect of age, $F < 1$, *ns*, but did reveal a main effect of feedback, $F(1,84) = 192.268$, $p < .0001$, $\eta_p^2 = .696$, and a significant interaction, $F(1,84) = 8.635$, $p = .004$, $\eta_p^2 = .093$. Follow up contrasts revealed college participants retained more errors after receiving feedback than did preschool participants, $F(1,84) = 3.489$, $p = .065$, $d = .761$, whereas, preschool participants retained more errors when they did not receive feedback, $F(1,84) = 7.672$, $p = .007$, $d = 1.674$ (see Table 4 for means and standard errors of retained responses).[6]

Changing Responses

A 2 (Feedback: yes, no) X 2 (Age: college, preschool) repeated measures ANOVA on the proportion of error correction revealed no main effect of age, $F < 1$, *ns*, but did reveal a main effect of feedback, $F(1,84) = 192.268$, $p < .0001$, $\eta_p^2 = .696$, and a

significant interaction between receiving feedback and age, $F(1,84) = 8.635$, $p = .004$, $\eta_p^2 = .093$. Follow up contrasts revealed preschool participants corrected more errors than did college participants after receiving feedback, $F(1,84) = 3.489$, $p = .065$, $d = .761$, whereas, preschool participants corrected fewer errors than did college participants when they did not receive feedback, $F(1,84) = 7.672$, $p = .007$, $d = 1.674$ (see Table 4 for means and standard errors of retained responses).

A 2 (Feedback: yes, no) X 2 (Age: college, preschool) repeated measures ANOVA on the proportion of new errors revealed a main effect of feedback, $F(1,104) = 10.391$, $p = .002$, $\eta_p^2 = .092$, an almost significant main effect of age, $F(1,104) = 3.249$, $p = .074$, $\eta_p^2 = .074$, and a significant interaction between feedback and age, $F(1,104) = 4.78$, $p = .034$, $\eta_p^2 = .044$. Follow up contrasts revealed no effect of feedback on the proportion of new errors for college students, $F < 1$, *ns*, whereas, there was an effect of feedback for preschool participants, such that they committed fewer new errors after receiving feedback than after not receiving feedback, $F(1,104) = 8.49$, $p = .006$, $d = 1.665$ (see Table 4 for means and standard errors of retained responses).

Test 1 and Test 2 Mean Confidence Rating

A 2 (Feedback, yes, no) X 2 (Test type: Test 1, Test 2) X 2 (Age: college, preschool) mixed ANOVA on mean confidence rating revealed a main effect of age, $F(1,104) = 4.005$, $p = .048$, $\eta_p^2 = .037$, a main effect of test type, $F(1,104) = 74.891$, $p < .0001$, $\eta_p^2 = .419$, and an almost significant effect of feedback, $F(1,104) = 2.88$, $p = .092$ $\eta_p^2 = .027$. The main effects were qualified by a significant interaction between feedback and test type, $F(1,104) = 23.176$, $p < .0001$, $\eta_p^2 = .182$, and an interaction that approached

significance between receiving feedback and age, $F(1,104) = 2.88$, $p = .092$, $\eta_p^2 = .027$. There was no interaction between test type and age, $F(1,104) = 2.708$, $p = .103$, $\eta_p^2 = .025$, and no three-way interaction between test type, age, and feedback, $F < 1$, *ns*. Follow up contrasts to interpret the interaction between feedback and test type revealed on Test 1 participants rated no feedback questions slightly higher than they rated feedback questions, $F(1,104) = 3.478$, $p = .065$, $d = .682$; whereas, on Test 2 participants rated no feedback questions lower than feedback questions, $F(1,104) = 20.794$, $d = 4.07$, $p < .0001$ (see Figure 12).

Correct Retention as a Function of Feedback and Confidence.

Because of the small sample sizes, examining correct retention and error correction as a function of confidence rating was analyzed separately for college participants and preschool participants.

Preschool participants. A 2 (Feedback: yes, no) X 2 (Confidence rating: high, low) repeated measures ANOVA on the retention of correct responses revealed no main effect of receiving feedback, $F(1,6) = 1.041$, $p = .347$, $\eta_p^2 = .148$, no main effect of confidence rating, $F(1,6) = 2.204$, $p = .188$, $\eta_p^2 = .269$, and no significant interaction, $F(1,6) = 2.204$, $p = .188$, $\eta_p^2 = .269$ (see Figure 13).

College participants. A 2 (Feedback: yes, no) X 2 (Confidence rating: low, high) repeated measures ANOVA on correct retention revealed a significant main effect of receiving feedback, $F(1,20) = 144.500$, $p < .0001$, $\eta_p^2 = .878$, a significant main effect of confidence, $F(1,20) = 107$, $p < .0001$, $\eta_p^2 = .843$, and a significant interaction between receiving feedback and confidence, $F(1,20) = 156.900$, $p < .0001$, $\eta_p^2 = .887$ (see Figure

14). Follow up contrasts revealed no effect of feedback on the retention of correct responses given in high confidence (rating of 3) on Test 1, $F(1,20) = 1.949$, $p = .178$, $d = .876$. But, when participants answered a response correctly in low confidence (rating of 1), they were more likely to also answer it correctly on Test 2 if they received feedback on the response than if they did not, $F(1,20) = 697.174$, $p < .0001$, $d = .660$.

Error Correction as a Function of Feedback and Confidence.

Preschool participants. A 2 (Feedback: yes, no) X 2 (Confidence rating: high, low) repeated measures ANOVA on error correction revealed a significant main effect of receiving feedback, $F(1,12) = 32.241$, $p < .0001$, $\eta_p^2 = .729$, but no main effect of confidence rating, $F < 1$, *ns*, nor an interaction between receiving feedback and error correction, $F(1,12) = 1.220$, $p = .291$, $\eta_p^2 = .092$ (see Figure 15). To test for the hypercorrection with feedback effect, an a priori repeated measures t-test examining feedback questions was conducted, but it revealed that participants did not correct more errors committed in high confidence ($M = .907$, $SE = .047$) than in low confidence after receiving feedback ($M = .917$, $SE = .061$), $t(17) = .169$, $p = .868$, $d = .081$.

College participants. A repeated measures t-test revealed participants corrected marginally more errors after receiving feedback than not receiving feedback for responses given in high confidence, $t(7) = 2.000$, $p = .087$, $d = .700$. Participants also corrected more low confident errors made after receiving feedback, $t(46) = 2.896$, $p = .011$, $d = .722$ (see Figure 16). A priori t-test also revealed there was no hypercorrection effect, as participants did not correct more errors committed in high confidence ($M =$

.769, *SE* = .122) than in low confidence after receiving feedback (*M* = .615, *SE = .128)*, *t*(12) = .843, *p* = .416, *d* = .486.

Discussion

Results provide compelling evidence that children benefit from receiving feedback after answering questions, and that this benefit persists at least ten minutes. Despite information processing limitations (e.g., Bjorklund & Harnishfeger, 1987) that could have diluted the impact of receiving feedback, children performed better on Test 2 after receiving feedback. Their improved performance stems largely from participants correcting most of their Test 1 errors, and to a lesser extent, participants retaining correct original responses from Test 1. This improvement supports the idea that receiving feedback would improve memory performance in children. These findings are also consistent with the hindsight bias research that demonstrates children are influenced by external information (e.g., Bernstein et al., 2007). Finally, from the source monitoring perspective, children benefit from receiving feedback because they are able to maintain a memory of the feedback and the original memory, and successfully discriminate between the two (e.g., Foley et al., 1983).

The college participants also benefited from receiving feedback, but their improved performance stemmed mainly from correcting Test 1 errors and not from retaining Test 1 correct responses. Please note that the college participants performed near ceiling on Test 1, making it unlikely to observe feedback enhancing the retention of correct responses. These findings indicate that providing feedback can improve memory performance for young children as well as college students. Because of the different

testing procedures, though, comparisons across the age groups should be made with caution.

The college participants demonstrated the same patterns in their confidence for the children's story materials as they did with the more difficult journal articles. There is an effect of receiving feedback on the retention of correct responses, but it is only evident for low confident correct responses and not high confident correct responses. If the confidence rating scale in Experiments 1 and 2 matched the scale in Experiment 3, college participants may have demonstrated greater confidence in responses in Experiment 3 because their performance was better. It is impossible to examine these differences because the scales are not equivalent. The preschool participants' confidence pattern revealed the same basic patterns as the college students. They were more confident after receiving feedback and more confident on Test 2 than Test 1. Feedback did not enhance the retention of correct responses for low confident correct responses as it did for the college participants. Because children demonstrate poorer monitoring abilities, they may be less of a relationship between feedback and confidence.

Despite the substantial timing differences in the procedures for the children and college students, comparing performance as a function of age can illustrate whether they demonstrated similar patterns of performance. Preschool and college participants demonstrated very similar patterns of performance. Most importantly, they both improved from Test 1 to Test 2 after receiving feedback because they corrected the majority of their errors made on Test 1. College participants started out with superior performance, but the preschool participants benefited slightly more from receiving

feedback (i.e., corrected more errors and retained more correct responses) than the college participants. This is impressive because the college students were performing the task on materials designed to be much easier for them, are more familiar with a structured learning situations, and have slower rates of forgetting (Howe, 1991). This similar pattern of performance may suggest feedback improves memory similarly for preschool children and college students, and that feedback may be used to improve memory performance in similar contexts. These results imply that educators for both young children and college students may see a benefit in student performance by including feedback as one tool in teaching.

CHAPTER IV

GENERAL DISCUSSION

Although memory performance is often accurate, it is now well established that forgetting and memory distortions commonly occur (Anderson & Neely, 1996; Koriat, Goldsmith, & Pansky, 2000; Roediger & McDermott, 2000). Because of these memory errors, it is important to discover methods of correcting memory mistakes and improving memory performance. In the field of cognitive psychology, the last several years has seen a resurgence of interest in education relevant cognitive phenomena, and much of that interest has focused on feedback (Butler et al., 2009; McConnell & Hunt, 2007; Pashler et al., 2005). Previous work has established that college students' test performance often improves from receiving feedback (e.g., Anderson et al., 1972; Bangert-Drowns et al., 1991; Kulhavy, White, Topp, Chan, & Adams, 1985; Kulik & Kulik, 1988). Several studies suggest that the confidence participants have in their original responses influences the effectiveness of the feedback (e.g., Bulter et al., 2008; Metcalfe & Butterfield, 2001). The purpose of the current experiments was to examine the relationship between monitoring performance and feedback, and to determine whether feedback improves memory and monitoring performance for college and preschool participants. Generally, the results from the experiments demonstrate that both preschool and college participants use feedback to influence their test performance. When feedback provided true information it improved their performance, and when it provided false information in

impaired their performance. Very few participants noticed the feedback included false information, but even those who did continued to use the false feedback to influence their Test 2 performance.

The findings from the three experiments all suggest that participants' confidence influences the effectiveness of feedback. More specifically, when college participants received true feedback they retained more correct responses when not very confident in their responses, participants were more confident after receiving feedback than after not, and receiving feedback improved participants' monitoring of their performance. The findings from Experiment 1 extended Butler et al.'s (2008) finding that feedback enhances the retention of correct responses. This is an important finding because determining whether receiving feedback can enhance the retention of correct responses has proved elusive for many years (Kulhavy et al., 1990; McConnell & Hunt, 2007; Pashler et al., 2005). In addition to demonstrating confidence is important for the effectiveness of feedback, these results suggest receiving feedback enhances participants' memory and monitoring performance.

The finding from Experiment 2 suggest that participants' confidence in the feedback may not prove as important as is their confidence in their original responses. More specifically, comparing participants' Test 2 performance in Experiment 1 and Experiment 2 demonstrates the influence of receiving true and false feedback on test performance. When participants received 100% true feedback their performance improved. This improved performance stems largely from increased error correction, but participants were also more likely to retain low confident correct responses. This benefit

supports previous findings that conclude receiving feedback improves test performance (e.g., Kulhavy & Anderson, 1972; McConnell & Hunt, 2007; Pashler et al., 2005). When participants received feedback that contained 50% false information, the benefits of receiving feedback dropped dramatically. When participants received false feedback they no longer improved on Test 2, and more precisely, they did not correct Test 1 errors and did not retain Test 1 correct responses. When comparing directly whether participants were more likely to use the true or the false feedback, results revealed participants were more likely to use the feedback that confirmed the original response. When participants answered Test 1 questions correctly they were more likely to follow the true feedback indicating they should retain their response; however, when they answered Test 1 questions incorrectly they were more likely to follow the false feedback indicating they should retain their response. These results may reflect a confirmation bias that participants are more accepting of information that confirms their beliefs as opposed to contradicts it (e.g., Baron, 2000; McKenzie, 2006).

Comparing participants' confidence in their responses in Experiment 1 and Experiment 2 demonstrates the influence of feedback on confidence ratings and indirectly indicates whether participants were aware of the false feedback. The confidence ratings indicate participants did not detect the false feedback, as participants provided similar confidence ratings for answers that received true feedback and false feedback, their Test 2 confidence ratings did not drop compared to Test 1, and their ratings were similar to participants who only received true feedback. Previous research reveals that participants' confidence in their Test 1 responses can moderate the effects of receiving feedback (e.g.,

Butler et al., 2008; Butterfield & Metcalfe, 2001). Accordingly, in these experiments after receiving true feedback participants had higher retention of Test 1 correct responses after receiving feedback, but it only occurred when their Test 1 correct responses were made in low confidence. These low confident correct responses can be viewed as guesses, and receiving confirmatory feedback confirms these correct guesses. Confidence in Test 1 *errors* did not moderate the effects of feedback, such that participants were no more likely to correct errors made in high confidence than errors made in low confidence. As mentioned previously, the lack of a hypercorrection with feedback effect may be due to the amount of time participants had with the feedback.

Using a more direct approach to answering whether participants detect and continue to use false feedback was to ask participants whether they noticed anything about the feedback. Fifteen participants (16% of the participants) reported being aware the feedback was false, and these participants had better performance than participants who reported being unaware. Aware participants' improved performance did not stem from being less likely to use the false feedback, but from aware participants answering more questions correctly on Test 1 than Test 2 and being more likely to follow the true feedback. These results suggest participants may have had difficulty discriminating between Test 1 responses and the feedback when completing Test 2. Because the aware participants noticed some of the feedback was false, their performance would have benefited from ignoring all of the feedback and focusing on their original Test 1 responses. Accordingly, several participants indicated they attempted to use this strategy during Test 2, but their performance suggests difficultly discriminating Test 1 responses

from the feedback, which supports using a source monitoring framework to understand how feedback works.

Source Monitoring Framework

According to a source monitoring explanation, the feedback and the original response should be viewed as two sources of information. When completing Test 2, the original response and the feedback may come to mind, and participants will attempt to discriminate the origins of those memories. To answer Test 2 questions that were answered incorrectly on Test 1, participants must respond with the feedback and not the original response. When feedback impairs memory for the original responses, source monitoring is unnecessary to ensure a correct response. However, as evident from feedback research, impairment of the original response is unnecessary for correct responding on Test 2, as seen in many experiments when participants who received feedback also remember their original responses (e.g., Brosvic, et al., 2006; Clariana, Wagner, & Roher-Murphy, 2000; Dihoff et al., 2004; Peeck & Tillema, 1978-1979). Peeck and Tillema asked participants to not only select the correct response, but to also select their original response, and they found that participants who received feedback were as likely as participants who did not receive feedback to select the original response. They also found that participants who answered correctly on Test 2 were as likely to remember the original response as were participants who answered the question incorrectly on Test 2. It is in situations that the feedback does not impair memory for the original response that source monitoring becomes helpful for error correction because participants must discriminate between original responses and feedback during Test 2. Failing to discriminate source should not

necessarily worsen performance. For example, if, at test, participants remember a response but cannot discriminate the source, they may answer correctly—but only if the feedback comes to mind or they answered correctly originally. If both responses come to mind, however, participants must attribute a correct source and select a correct response.

In Experiment 1 all of the feedback was correct, so it represents a normal feedback situation. In this situation participants may answer correctly on Test 2 for different reasons. First, if the feedback impairs memory for the original response, they will answer correctly on Test 2 regardless of how they responded on Test 1. On the other hand, because previous research demonstrates consistently that feedback works without impairing memory for the original response, participants may use source monitoring to answer correctly on Test 2. If the feedback does not impair memory for the original response but participants successfully discriminate memory for the original response and the feedback, they will answer correctly regardless of how they answered on Test 1. When participants fail to discriminate between Test 1 responses and the feedback, participants may answer correctly if they happen to pick the feedback regardless of how they answered on Test 1. Yet, if they do not pick the feedback, they may still answer correctly if they pick their original response and they happened to answer it correctly on Test 1. They will answer incorrectly on Test 2 if they pick their original response and they happened to answer incorrectly on Test 1.

In Experiment 2, because 50% of the feedback was false, there were fewer possibilities to answer questions correctly. In this situation, if feedback blocked memory for the original response participants would answer questions correctly when they received true feedback and

answer questions incorrectly when they received false feedback. The results of Experiment 2 support this idea because participants did better after receiving true feedback than false feedback. However, because previous work demonstrates that it is not necessary for feedback to impair memory of the original response to be effective, a source monitoring perspective is appropriate to understand what may happen when feedback does not impair the original memory. If participants were unaware that some of the feedback was false, then they should follow the same path as in Experiment 1, but their performance would suffer. Specifically, if they discriminate between Test 1 responses and the feedback, they will answer correctly when they receive true feedback. If they do not discriminate the feedback from their original responses, they will answer correctly when they select the true feedback or when they select their original correct response. If they do not discriminate the feedback they will answer questions incorrectly when they select the false feedback or when they select their original incorrect response.

Participants who were aware some of the feedback was false should not follow the same path when discriminating between Test 1 responses and feedback because they would doubt the correctness of the feedback. If aware participants can distinguish the false feedback from the true feedback, when they discriminate the feedback from the original response they should use the feedback when it was true and use their original responses when the feedback was false. It would be surprising, however, if the participants would be able to distinguish between all the true and false feedback, and this would probably only happen if participants were very familiar with the material. If participants suspected that some of the feedback was false, but were unable to discriminate which was false, the aware participants would have better

Test 2 performance if they responded with their original responses and not the feedback. Because aware participants started out answering over 50% of Test 1 questions correctly, their Test 2 performance would fare better sticking to their original responses and ignoring the feedback.

Although several of the aware participants indicated they tried to ignore the feedback when answering Test 2 questions, they still used the true and the false feedback. Because research suggests feedback does not typically impair memory for the original responses, these participants seem to have had difficulty discriminating their Test 1 responses from the feedback. They answered Test 2 questions correctly when they answered with the true feedback or their Test 1 correct responses, but they answered Test 2 questions incorrectly when they answered with the false feedback or their original Test 1 incorrect responses. The results from Experiment 2 revealed that participants answered questions incorrectly on Test 2 largely because they were changing Test 1 correct responses to errors on Test 2 following the false feedback. Unfortunately only 15 participants reported being aware some of the feedback provided false information, and conclusions that participants continue to use the false feedback to influence performance, even after detecting it was false, should be made with caution. Future studies need to provide false feedback that more participants detect is false. With more participants detecting that some of the feedback is false, it will be easier to find any differences between aware and unaware participants. There are two potential methods to encourage more participants being aware of the false feedback. First, using materials more familiar to college participants should increase the likelihood participants will notice the false feedback. Although participants had 15 minutes to read each article, they

still answered many questions incorrectly on Test 1 suggesting unfamiliarity with the topic, which may have made it difficult to notice inconsistencies with the feedback. When participants are familiar with the topic they may have the outside knowledge to help detect inconsistencies in the material, and not have to rely solely on remembering information they read in a short study session. Second, increasing the percentage of questions that receive false feedback should increase the likelihood of participants noticing the false feedback. In Experiment 2, 50% of the feedback contained false answers, but if this percentage was increased more participants may be aware that some of the feedback is false because the feedback would contradict information they explicitly remember reading. If these situations increase the number of participants who detect the false feedback, a better test of whether participants continue to use the feedback when they detect it is false would be possible.

Strictly using a source monitoring framework does not account for all of the findings. In addition to viewing the original response and the feedback as two separate sources, participants' confidence in at least their original response is also important. For example, receiving feedback only increases the retention of correct responses when the responses are made in low confidence. Ignoring the participant's confidence in their Test 1 responses and simply examining whether there is an effect of feedback makes the effects of receiving feedback on correct responses less clear. In Experiment 2, participants still used the feedback even when they detected it was false. This may suggest that participants' confidence in the feedback is not as important as their confidence in their original responses. However, very few participants actually noticed

that the feedback was false, which makes it difficult to conclude that confidence in the feedback is unrelated to participants' choosing to use the feedback.

A source monitoring framework supports the conclusions about feedback found in other studies. Previous studies suggest that delayed feedback is more effective than immediate feedback (e.g., Peeck & Tillema, 1978-1979; Phye, 1979; Phye & Andre, 1989). According to a source monitoring framework, the greater the temporal separation between the feedback and the original response the more distinct the memory records because the context may change, resulting in improved source monitoring (e.g., Drummey, & Newcombe, 2002; Johnson et al., 1993). As the delay between the original responses and the feedback increases, the context cues should increase and discrimination becomes easier. Also, a source monitoring framework is compatible with the findings that participants can remember their original responses after receiving feedback. According to a source monitoring framework, the original response and the feedback exist as separate traces and the participant must attribute a source.

Developmental Origins of Feedback

Results from Experiment 3 revealed that preschool children benefit from receiving feedback. Impressively, the preschool participants corrected over 90% of the Test 1 errors, which was equivalent to the error correction rate for college participants using the same materials. Although the college participants completed the experiment using the same materials as the preschool children, the procedures differed, and the biggest procedural change between the college participants and the preschool participants was the delay between receiving the feedback and Test 2. The preschool participants completed Test 2 ten minutes after receiving

the feedback, while the college participants completed Test 2 forty-eight hours after receiving the feedback. Because preschool children's test performance improves from receiving feedback, it is important to understand more about the benefits of receiving feedback and the potential constraints. One potential area is to understand more about how long the benefits of receiving feedback persist. Specifically, a future study should increase the delay of Test 2 to time lapses more similar to college participants, and compare the effects of receiving feedback at different ages. This is important for several reasons. First, if children and college participants complete the experiment with more similar procedures, it will allow for a more conclusive comparison of the effects of feedback at different ages. This has both theoretical and applied importance. Theoretically, if children do not benefit from receiving feedback as compared to college participants, it may indicate that cognitive processes that are less developed in children are involved in how feedback improves memory performance. From an applied perspective, if children benefit less from receiving feedback than college participants, then researchers can focus on discovering what types of feedback may best benefit children or what other techniques to use. Increasing the delay between receiving feedback and completing Test 2 is also important to discover any constraints to the benefit of receiving feedback. Receiving feedback overwhelmingly improved memory performance for preschool children, and that benefit persists for at least 10 minutes. If receiving feedback only improves memory performance for such a brief time, however, providing feedback may not be a useful tool for improving children's long term memory performance.

Despite the differences, the fact that the preschool children benefited from receiving feedback at a similar level as the college participants, demonstrates clearly that children can benefit from receiving feedback. It is impressive that children showed similar benefits as the college participants despite their limitations in their cognitive system, and despite that the college participants completed the test on materials much easier for them than the preschool participants. In fact, when the college participants completed the age appropriate task in Experiment 1, they show less of a benefit from feedback than did the preschool children on their age appropriate task in Experiment 3.

Children may have performed better than the college participants for several reasons. First, the college participants' enhanced delay between receiving feedback and completing Test 2 may have hurt their performance because they experienced more difficulty forgetting the information (Ebbinghaus (1885/1964), or because they had more difficulty discriminating the original response and the feedback. Several findings demonstrate that participants have more difficulty discriminating source as the delay between the sources increased (Bornstein & Lecompte, 1995; Frost, Ingraham, & Wilson, 2002). Because college participants had a longer delay, they may have had more difficulty discriminating responses that came to mind. Second, the college participants' metacognitive expectations of the situation may have worsened their performance. College participants' are overconfident when they are presented with an easy task. Moore and Nealy told participants they would be answering easy questions or difficult questions, and before reading the questions participants judged how many questions they would answer correctly. Participants overestimated their performance answering the easy

questions. When participants are overconfident they exert less effort in the task. Because the college participants were completing both the children's materials and the more difficult journal articles, they have underestimated the effort needed to remember the story content. Related, findings examining children's strategy use demonstrates that when children exert more effort than college students, they compensate for any initial memory differences (Bjorklund & Harnishfeger, 1987).

A source monitoring framework can also account for how young children benefit from feedback. Research examining children's ability to monitor sources suggests children can distinguish between two external sources (e.g., Drummey & Newcombe, 2002; Foley et al., 1983; Lindsay, Johnson, & Kwon, 1991; Powell & Thomson, 1997). This suggests that if the Test 1 response and the feedback come to mind, children may be able to discriminate the Test 1 response from the feedback, resulting in a benefit from receiving feedback as seen in Experiment 3. Previous findings demonstrate, however, that children's source monitoring is impaired more from very similar sources than is adults' source monitoring. Lindsay et al. (1991) found that children successfully discriminate dissimilar external sources but have difficulty discriminating very similar sources. In their experiment, children and adults listened to words projected through either the left or the right side of a headset. These words were either presented with the same gender on both sides of the headset, or one side of the headset was always associated with one gender. They found that children and adults successfully discriminate source when the voices were different genders (dissimilar sources), but only children had problems discriminating the source when the voices were the same gender (similar sources). These results suggest children's performance would suffer to a greater extent than

adults after manipulations that increase the similarity between Test 1 responses and the feedback.

One potential method for testing the appropriateness of a source monitoring framework for both ages is to combine source monitoring procedures and feedback procedures into one task, and then examine if participants who do well with the source monitoring procedures also do well with the feedback procedures. For example, during Test 2 participants could be asked to respond with all answers that come to mind, regardless if they were correct. The participants would then attribute the source of the memory to each answer and then select the correct answer. Participants who did better with source judgments should also answer more Test 2 questions correctly. These methods could also be separated into different tasks (i.e., a source monitoring task and a feedback task), and participants' performance on the two separate tasks can be examined to determine if there were correlations between how well one performed in a source monitoring task and how well one performed in a feedback task.

Metacognition

Participants demonstrated relative accuracy in their monitoring as they gave higher confidence ratings to questions they answered correctly than questions they answered incorrectly. In Experiment 1, when participants received feedback their monitoring improved, suggesting that receiving outside information about their performance improved their monitoring. This finding confirms other research demonstrating participants use cues when judging their performance (e.g., Hertzog, Kidder, Powell-Moman, & Dunlosky, 2002; Koriat, 1997). As mentioned previously, Koriat (1997) proposed the cue-utilization approach and found participants use different cues when predicting how well they have learned material.

Similarly, Gonzalez-Vallejo and Bonham (2007) found that if participants are rewarded or punished after providing confidence judgments, their monitoring becomes more accurate. Receiving feedback about Test 1 would be a strong cue about one's performance that participants may continue to use when assessing Test 2 performance. In Experiment 2, participants' monitoring suffered from receiving false feedback, such that participants were not giving Test 2 correct responses higher confidence ratings than Test 2 incorrect responses. This may suggest participants used the feedback when making judgments, but misinformation impaired the process. These findings support Koriat and Goldsmith's (1996) claim that participants can successfully monitor under some situations but not others. They termed this relative accuracy monitoring resolution. They found that participants demonstrated monitoring resolution when they answered normal trivia questions, but when participants answered tricky trivia questions their monitoring resolution suffered. The findings from the current studies demonstrate that receiving feedback improves participants' ability to monitor their performance.

Participants used the feedback when it provided true and false information. This implies participants were influenced by the feedback, regardless of the correctness of the information. The vast majority of participants did not detect the feedback was false, even after providing highly confident ratings and answering correctly, which suggests participants may trust the feedback more than their own memories. The preschool participants also used the feedback, and may have been even more likely to use it than the college participants, as they corrected more errors after receiving feedback than did the college participants. Previous research suggests children are prone to relying on the memories or suggestions from other

sources (e.g., Bjorklund, Bjorklund, Douglas, & Cassel, 1998; Ceci, Ross, & Toglia, 1987). Ceci et al. found that young children changed their correct memories when an authority figure repeatedly gave misleading information. These findings suggest if children were given false feedback they also would have been likely to respond with the incorrect information.

Future Directions

These results from the experiments demonstrate that participant use true feedback to influence performance positively, use false feedback to influence performance negatively, and that the preschool and college participants both benefit from receiving feedback. There are 3 separate programs of research that may improve understanding of how and when feedback is effective. They are: (a) developing a theoretical explanation of feedback, (b) providing a better understanding of when feedback will improve children's memory performance, and (c) developing an applied focus to feedback research. Although a source monitoring framework has been proposed as one possible method for understanding how feedback works, several future steps are necessary to understand more about feedback. From a source monitoring perspective, participants should be making decisions about the source of their memories by judging the quality of the memory and then attributing it to a source. If feedback influences memory through source monitoring processes, then during Test 2 participants should be judging the memories that come to mind and attributing it to the feedback or original test. Because these experiments were designed to test whether participants use false feedback and whether they detect false feedback, they do not test the appropriateness or limits of a using a source monitoring approach to understanding how receiving feedback influences Test 2 performance. A source monitoring framework was proposed because post hoc explanations of

the effects of feedback (e.g., benefit of delayed feedback, memory for original responses, difficulty discriminating false feedback) fit within a source monitoring framework. However, the next step in assessing if a source monitoring framework is appropriate to understand how feedback works is to design a program of research that makes and tests predictions based on source monitoring. If a source monitoring framework is appropriate, one should expect manipulations that influence source monitoring to influence feedback's effectiveness. Testing these assumptions will demonstrate whether a source monitoring framework is an appropriate account of feedback. If a source monitoring framework proves an unsuccessful explanation of how feedback influences memory performance, other unexplored theoretical approaches need to be developed and examined.

Both the preschool children and the college participants benefited from receiving feedback, but further steps are needed to understand more about the benefits of receiving feedback for young children. In addition to creating situations that allow for better comparisons between children and college students, a program of research should examine other conditions in which children can receive feedback. For example, future research needs to examine the learning situation in which children receive feedback. In Experiment 3, the children received feedback on a fictional story, but it is unclear if this benefit would extend to more complex materials similar to what they learn in school. Receiving feedback may be more important for learning in complex situations where the children have difficulty with the material. It is in these situations that feedback may be less effective because their cognitive resources may be needed to complete the difficult task, and processing and integrating the feedback may diminish those resources (e.g., Bjorklund & Harnishfeger, 1987).

The findings from these experiments and others (e.g., Anderson et al., 1972; Andre & Theiman, 1988; Butler et al., 2008; Kulhavy & Anderson, 1972; McConnell & Hunt, 2007) demonstrate consistently that participant's performance improves from receiving feedback, but many of the studies examining feedback used laboratory situations that differ sharply from typical classroom environments (e.g., Anderson et al., 2007; Kulhavy & Anderson, 1972; McConnell & Hunt, 2007). Although not within the scope of these experiments, a future line of research needs to examine the benefits of receiving feedback either in classrooms or in situations that more closely mimic classroom learning. Examining several critical differences between typical feedback experiments and typical classrooms would be a good starting point for a more ecologically valid line of research. These key differences include: (a) students' familiarity with the material, (b) the delays between studying the material and receiving a test on the material, (c) delays between receiving the feedback and completing Test 2, and (d) the availability of the feedback and study material during the delays. Examining whether feedback continues to improve performance under these conditions will provide a clearer picture of the limits and applications of receiving feedback.

In conclusion, the results from these experiments provided new information about the relationship between monitoring and feedback. Specifically, participants' confidence is important in examining the effects of feedback. In Experiment 1, participants benefited from receiving feedback, and their confidence in the original responses moderated the effects. When participants completed a difficult test that encouraged low confident correct responses, receiving feedback enhanced the retention of correct responses. Experiment 2 was the first study to provide false feedback without a warning, and to examine the effects of false feedback

on participants' performance and confidence ratings. These results suggest participants are easily influenced from feedback, even when the feedback contains misinformation. Experiment 3 was the first experiment examining the effects of receiving feedback for both preschool and college participants, and the results demonstrated that both children and adults benefit from receiving feedback. In general, the current experiments reveal that participants use feedback to influence their memory performance and their assessment of their performance. Participants use the feedback regardless if it contains incorrect information and regardless if they notice it contains incorrect information. The feedback and the original memory can be viewed as two separate sources of information, and this may provide a framework for understanding more about feedback and designing future studies.

FOOTNOTES

[1]In another experiment, that uses part of the same subject pool, 80% females and 30% minorities participated in the experiment.

[2]All follow up contrasts were linear contrasts where the coefficients of the groups sum to zero.

[3]These groupings were unequal because more responses were rated as 70 - 100 percent confident than the lower level of confidence ratings, so including a larger range for the lower confidence ratings allows for more similar amounts of responses in the lower and higher confidence rating groupings. All the analyses were also conducted using 100 – 70 for high and 29 – 0 for low and the same patterns emerge.

[4]All a priori contrasts were planned linear contrasts where the coefficients of the groups sum to zero. These linear contrasts were conducted because of reasons to expect differences between true feedback, no feedback, and false feedback.

[5]The single sample t-test included Test 1 participants that did not have Test 2 confidence ratings.

[6]*One participant did not make an error on Test 1.*

REFERENCES

Abeles, P., & Morton, J. (1989). Avoiding misinformation: Reinstating target modality. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 52A*, 581–592.

Anastasio, P.A., Rose, K.C., & Chapman, J. (1999). Can the media create public opinion? A social identity approach. *Current Directions in Psychological Science, 8*, 152-155.

Anderson, M.C., & Neely, J.H. (1996). Interference and inhibition in memory retrieval. In: E.L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 237-313). San Diego, CA: Academic Press.

Anderson, R.C., Kulhavy, R.W., & Andre, T. (1972). Conditions under which feedback facilitates learning from programmed lessons. *Journal of Educational Psychology, 63, 186-188.*

Andre, T., & Theiman, A. (1988). Level of adjunct question, type of feedback, and learning concepts by reading. *Contemporary Educational Psychology, 13*, 296-307.

Anastasio, P.A., Rose, K.C., & Chapman, J. (1999). Can the media create public opinion? A social identity approach. *Current Directions in Psychological Science, 8*, 152-155.

Bahrick, H.P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General, 113*, 1-29.

Bangert-Drowns, R.L., Kulik, C.L.C., Kulik, J.A., & Morgan, M.T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*, 213-238.

Bernstein, D.M., Atance, C., Metlzoff, A.N., & Loftus, G.R. (2007). Hindsight bias and developing theories of mind. *Child Development, 78,* 1374–1398.

Birch, S.A.J., & Bernstein, D.M. (2007). What can children tell us about hindsight bias: A fundamental constraint on perspective-taking? *Social Cognition, 35*, 98-113.

Bisanz, G.L., Vesonder, G.T., & Voss, J.F. (1978). Knowledge of one's own responding and the relation of such knowledge to learning: A developmental study. *Journal of Child Experimental Psychology, 25*, 116-128.

Baron, J. (2000). *Thinking and deciding* (3rd ed.). New York: Cambridge University Press.

Bjork, R.A. (1988). Retrieval practice and the maintenance of knowledge. In M. Gruneberg & P. Sykes (Eds). *Practical Aspects of Memory: Current Research and Issues, Vol 1: Memory and Everyday Life* (396-401). Oxford, John Wiley & Sons.

Bjorklund, D.F., Bjorklund, B.R., Douglas Brown, R., & Cassel, W.S. (1998). Children's susceptibility to repeated questions: How misinformation changes children's answers and their minds. *Applied Developmental Science,* 2, 99-111.

Bjorklund, D.F., & Harnishfeger, K. (1987). Developmental differences in the mental effort requirements for the use of an organizational strategy in free recall. *Journal of Experimental Child Psychology, 44*, 109-125.

Brainerd, C.J., & Reyna, V. F. (1990). Inclusion illusions: Fuzzy-trace theory and perceptual salience effects in cognitive development. *Developmental Review, 10*, 365-403.

Brainerd, C.J., & Reyna, V.F. (1995). Learning rate, learning opportunities, and the development of forgetting. *Developmental Psychology, 31*, 251-262.

Brainerd, C.J. & Reyna, V.F. (2004). Fuzzy-trace theory and memory development. *Developmental Review*, 24, 396-439.

Brosvic, G.M., Epstein, M.L., Dihoff, R.E., & Cook, M.J. (2006). Acquisition and retention of esperanto: The case for error correction and immediate feedback. *The Psychological Record*, *56*, 205-218.

Butler, A.C., Karpicke, J.D., & Roediger, H.L. III. (2008). Correcting a metacognitive error: Feedback increases the retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 918-928.

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27, 1491-1494.*

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition Learning*, 1, 69-84.

Ceci, S.J., Ross, D.F., Toglia, M.P. (1987). Suggestibility of children's memory: Psycholegal implications. *Journal of Experimental Psychology: General, 116*, 38-49.

Cinan, S. (2006). Age-related changes in concept formation, rule switching, and perseverative

behaviors: A study using WCST with 12 unidimensional target cards. *Cognitive Development, 21*, 377-382.

Clariana, R.B., Ross, S.M., Morrison, G.R. (1991). The effects of different feedback strategies using computer-administered multiple-choice questions as instruction. *Educational, Technology, Research, and Development, 39, 5-17.*

Clariana, R.B., & Wagner, D., & Roher-Murphy, L.C. (2000). Applying a connectionist description of feedback timing. *Educational Technology, Research, and Development, 48, 4-21.*

Connor, L.T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging*, *12*, 50-71.

Crone, E.A., Ridderinkhof, K., Worm, M., Somsen, R.J.M., & van der Molen, M.W. (2004). Switching between spatial stimulus-response mappings: A developmental study of cognitive flexibility. *Developmental Science, 7*, 443-455.

Crone, E.A., Somsen, R.J.M., Zanolie, K., & van der Molen, M.W. (2006). A heart rate analysis of developmental change in feedback processing and rule shifting from childhood to early adulthood. *Journal of Experimental Child Psychology*, *95*, 99-116.

Deese, J. (1959). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports*, *5*, 305-312.

Denney, N. (1974). Evidence for developmental changes in categorization criteria for children and adults. *Human Development, 17*, 41-53.

Dihoff, R.E., Brosvic, G.M., Epstein, M.L. (2003). The role of feedback during academic testing: The delay retention effect revisited. *The Psychological Record, 53*, 533-548.

Dihoff, R.E., Brosvic, G.M., Epstein, M.L., & Cook, M.J. (2004). Provision of feedback during preparation for academic testing: Learning is enhanced by immediate but not delayed feedback. *The Psychological Record, 54,* 207-231.

Drummey, A.B., & Newcombe, N.S., (2002). Developmental changes in source memory. *Developmental Science, 5*, 502-513.

Dunlosky, J., Rawson, K.A., & Middleton, E.L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypothesis. *Journal of Memory and Language, 52*, 551-565.

Dunlosky, J., & Thiede, K.W. (1998). What makes people study more? An evaluation of factors that affect self-paced study. *Acta Psychologica, 98*, 37-56.

Eagly, A.H., Kulesa, P., & Chen, S. (2001). Do attitudes affect Memory? Tests of the congeniality hypothesis. *Current Directions in Psychological Science. 10*, 5-9.

Garry, M., & Polaschek, D.L.L. (2000). Imagination and memory. *Current Directions in Psychological Science, 9*, 6-10.

Gonzalez-Vallejo, C., & Bonham, A. (2007). Aligning confidence with accuracy: Revisiting the role of feedback. *Acta Psychologica, 125*, 221-239.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*, 404-411.

Ebbinghaus, H. (1885/1964). *Memory: A Contribution to Experimental Psychology*. Oxford England: Dover.

Erdfelder, E., & Buchner, A. (1998). Decomposing the hindsight bias: A multinomial processing tree model for separating recollection and reconstruction in hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 387-414.

Fischhoff, B. (1975). Hindsight does not equal foresight: The effect of outcome knowledge on judgment uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 288-299.

Fischhoff, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 349-358.

Fischhoff, B., & Beyth, R. (1975). “I knew it would happen" Remembered probabilities of once future things. *Organizational Behavior and Human Performance, 13*, 1-16.

Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906-911.

Hanaba, J.M., & Zaragoza, M.S. (2006). Interviewer feedback in repeated interviews involving forced confabulation. *Applied Cognitive Psychology, 21*, 438-455.

Hacker, D. (1998). Self-regulated comprehension during normal reading. In D. Hacker & J. Dunlosky (Eds.), *Metacognition in Educational Theory and Practice* (pp. 165-191). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Hart , J.T. (1965). Memory and the feeling of knowing experience. *Journal of Educational Psychology, 56*, 208-216.

Hart, J.T. (1966). Methodological note on feeling-of-knowing experiments. *Journal of Educational Psychology, 57*, 347-349.

Hart , J.T. (1967). Second-try recall, recognition, and the memory-monitoring process. *Journal of Educational Psychology, 58,* 193-197.

Hasher, L., Attig, M.S., & Alba, J.W. (1981). I knew it all along or did I. *Journal of Verbal Learning and Verbal Behavior, 20*, 86-96.

Hell, W., Gigerenzer, G., Gauggel, S., Mall, M., & Muller, M. (1988). Hindsight bias: An interaction of automative and motivational factors? *Memory and Cognition,* 16, 533-538.

Hertzog, C., Kidder, D.P., Powell-Norman, A., & Dunlosky, J. (2002). Aging and monitoring associative learning is monitoring accuracy spared or impaired? *Psychology and Aging, 17*, 209-225.

Hoff, S. (1958). *Danny and the Dinosaur*. New York: Harper Collins.

Howe, M.L. (1991). Misleading children's story recall: Forgetting and reminiscence of the facts. *Developmental Psychology, 27*, 746-762.

Hunt, R.R., & Ellis, H.C. (2004). *Fundamentals of Cognitive Psychology* (7th ed). New York: McGraw-Hill.

Hunt, R.R., Smith, R.E., & McConnell, M.D. (2009). When does feedback affect correct responses and why? *Manuscript being revised for resubmission.*

Jenkins, J.G., & Dallenbach, K.M. (1924) Obliviscence during sleep and waking.

American Journal of Psychology, 35, 605- 612.

Johnson, M.K. (2006). Memory and reality. *American Psychologist, 61*, 760-771.

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3-28.

Kang, S.H.K., McDermott, K.B., Roediger, H.L. III, (2007). Test format corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19*, 528-558.

Kirwan, W. (2007). *Nobody Notices Minerva.* New York: Sterling.

Koriat, A. (1993). How do we know what we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*, 609-639.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349-370.

Koriat, A. (2008). Subjective confidence in one's answers: The consensuality principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 945-959.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490-517.

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology, 51*, 481-537.

Koriat, A., Lichtenstein, & Fischhoff, B. (1980). Reasons for Confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 107-118.

Kulhavy, R.W. (1977). Feedback in written instruction. *Review of Educational Research, 47,* 211-232.

Kulhavy, R.W., & Anderson, R.C. (1972). Delay-retention effect with multiple-choice tests. *Journal of Educational Psychology, 63, 505-512.*

Kulhavy, R.W., & Stock, W.A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1*, 279-308.

Kulhavy, R.W., Stock, W.A., Hancock, T.E., Swidell, L.K., & Hammrich, T.E. (1990). Written feedback: Response certitude and durability, *Contemporary Educational Psychology, 15*, 319-332.

Kulhavy, R.W., Yekovich, F.R., & Dyer, J.W. (1976). Feedback and response confidence. *Journal of Educational Psychology, 68*, 522-528.

Kulhavy, R.W., Yekovich, FR., & Dyer, J.W. (1979). Feedback and content review in programmed instruction. *Contemporary Educational Psychology, 4*, 91-98.

Kulig, J.W. & Tighe, T.J. (1976). Subproblem analysis of discrimination-learning: Stimulus choice and response latency. *Bulletin of the Psychonomic Society, 7*, 377-380.

Kulik, J.A, & Kulik, C.L.C. (1988). Timing of feedback and verbal learning. *Review of Educational Research, 58*, 79-97.

Lindsay S.D. (1993). Eyewitness suggestibility. *Current Directions in Psychological Science*, *2*, 86-89.

Lindsay, D.S., & Johnson, M.K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition, 17*, 349-358.

Lindsay, D.S., & Johnson, M.K. (1989). The reversed eyewitness suggestibility effect. *Bulletin of Psychonomic Society, 27*, 111-113.

Lindsay, S.D., Johnson, M.K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology, 52*, 297-318.

Loftus, E.F. (1979). Reactions to blatantly contradictory information. *Memory & Cognition, 7*, 368-374.

Loftus, E.F., Miller, D.G., & Burns, H.J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19-31.

Marsh, E.J., Roediger, H.L. III, Bjork, R.A., & Bjork, E.L. (2007). The memorial consequences of multiple-choice testing. *Psychonomic Bulletin & Review,* 14, 194-199.

Marcovitch, S., & Zelazo, P.D. (2009). A hierarchical competing systems model of the emergence and early development of executive function. *Developmental Science, 12*, 1-25.

Markowitz , N., & Renner, K.E., (1966). Feedback and the delay-retention effect. *Journal of Experimental Psychology*, *72*, 452-455.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 114*, 1-16.

McConnell, M.D. (2007). A review and extension of the feedback literature: Offering

explanations for how feedback influences memory performance. Unpublished preliminary examination, University of North Carolina at Greensboro.

McConnell, M.D., & Hunt, R.R. (In press). Can false memories, once expressed, be corrected by feedback in the DRM paradigm? The effect of receiving feedback on critical errors in the DRM paradigm. *Memory & Cognition,* 35, 999-1006.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 114*, 1-16.

McDermott, K.B. (2006). Paradoxical effects of testing: Repeated retrieval attempts enhance the likelihood of later accurate and false recall. *Memory & Cognition, 34*, 261-267.

McGeoch, J.A. (1942). *The Psychology of Human Learning*. Oxford: Longmans Green.

McKenzie, C.R.M. (2006). Increased sensitivity to differentially diagnostic answers using familiar materials: Implications for confirmation bias. *Memory & Cognition, 34*, 577-588.

Melton, A.W., & Irwin, J.M, (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology, 53*, 173-203.

Metcalfe, J., Schwartz, B.L., & Joaquim, S.G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 851-861.

Miller, P.H., Seier, W.L., Barron, K.L., & Probert, J.S. (1994). What causes a memory strategy utilization deficiency? *Cognitive Development, 9*, 77-101.

Moore, D.A., & Healy, P.J. (2008). The trouble with overconfidence. *Psychological Review,* 115, 502-517.

Nelson, T.O., & Dunlosky, J. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language*, *33*, 545-565.

Nelson, T.O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 26, pp. 125-173). New York: Academic Press.

Pashler, H., Cepeda, N.J., Wixted, J.T., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 3-8.*

Payne, D.G., Toglia, M.P., Anastasi, J.S. (1994). Recognition performance level and the magnitude of the misinformation effect in eyewitness memory. *Psychonomic Bulletin & Review,* 1, 376-382.

Peeck, J., & Tillema, H.H. (1978-1979). Delay of feedback and retention of correct and incorrect responses. *Journal of Experimental Education, 47*, 171-178.

Peeck, J., Van Den Bosch, V.D., & Kreupeling, W.J. (1985). Effects of informative feedback in relation to retention of initial responses. *Contemporary Educational Psychology, 10*, 303-313.

Phye, G.D. (1979). The processing of informative feedback about multiple-choice test performance. *Contemporary Educational Psychology, 4*, 381-394.

Phye, G.D., & Andre, T. (1989). Delayed retention effect: Attention, perseveration, or both? *Contemporary Educational Psychology, 14,* 173-185.

Phye, G. D., & Baller, W. (1970). Verbal retention as a function of the informativeness and delay of informative feedback: A replication. *Journal of Educational Psychology, 61*, 380-381.

Pohl, R.F. (1998). The effects of feedback source and plausibility of hindsight bias. *European Journal of Cognitive Psychology*, *10*, 191-212.

Pohl, R.F., & Hell, W. (1996). No reduction in hindsight bias after complete information and repeated test. *Organizational Behavior and Human Decision Processes, 67*, 49-58.

Powell, M.B., & Thomson, D.M. (1997). Contrasting memory for temporal-source and memory for content in children's discrimination of repeated events. *Applied Cognitive Science, 11*, 339-360.

Pressley, M., Levin, J.R., & Ghatala, E.S. (1984). Memory strategy monitoring in adults and children. *Journal of Verbal Learning and Verbal Behavior, 23*, 270-288.

Roebers, C.M. (2002). Confidence judgments in children's and adults' event recall and suggestibility. *Developmental Psychology, 38*, 1052-1067.

Rantzen, A. & Markham, R. (1992). The reserved eyewitness testimony design: More evidence for source monitoring. *Journal of General Psychology*, *119*, 37-43.

Raye, C.L., Johnson, M.K. & Taylor, H. (1980). Is there something special about

memory for internally generated information? *Memory & Cognition, 8*, 141-148.

Roediger, H.L., & Karpicke, J.D. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language,* 57, 151-162.

Roediger, H.L. III, & McDermott, K.B. (2000). Distortions of memory. In E. Tulving & F.I.M Craik (Eds.). *The Oxford Handbook of Memory* (pp. 149-162).

Schneider , W., & Bjorklund, D.F. (1998). Memory. In D. Kuhn & R.S. Siegler (Vols. Eds.), *Cognitive, Language, and Perceptual Development*, Vol 2. In W. Damon (Gen. Ed.), *Handbook of Child Psychology.* New York: Wiley.

Serra, M.J., & Dunlosky, J. (2005). Does retrieval fluency contribute to the underconfidence-with-practice effect? *Journal of Experimental: Learning, Memory, and Cognition, 31*, 1258-1266.

Son, L.K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 204-221.

Stock, W.A., Kulhavy, R.W., Pridemore, D.R., & Krug, D. (1992). Responding to feedback after multiple-choice answers: The influence of response confidence. *The Quarterly Journal of Experimental Psychology, 45A*, 649-667.

Sturges, P.T. (1972). Informative delay and retention: Effect of information in feedback and tests. *Journal of Educational Psychology, 63*, 32-43.

Templeton, L.M., & Wilcox, S.A. (2000). A tale of two representations: The misinformation effect and children's developing theory of mind. *Child Development, 71*, 402-416.

Treiman, R. (2000). The foundations of literacy. *Current Directions in Psychological Science, 9*, 89-92.

Troyer, A.K., Winocur, G., Craik, F.I.M., & Moscovitch, M. (1999). Source memory and divided attention: Reciprocal costs to primary and secondary tasks. *Neuropsychology, 13*, 467-474.

Tulving, E. (1967). The effects of presentation and recall of material in free-recall learning. *Journal of Verbal Learning and Verbal Behavior,* 6, 175-184.

Zelazo, P.D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences*, *8*, 12-17.

APPENDIX A. TABLES

Table 1.
Experiment 1 Mean Proportion of Retention of Responses and Changing Responses.

	Feedback		No Feedback	
Proportion	Mean	SE	Mean	SE
Retention of Correct Responses	.739	.017	.690	018
Retention of Errors	.494	.016	.748	.016
Error Correction	.506	.018	.252	.016
New Errors	.261	.018	.311	.018

Table 2.
Experiment 2 Mean Proportion of Retention of Responses and Changing Responses as a Function of Feedback Type.

	Feedback					
	True Feedback		No Feedback		False Feedback	
Proportion	Mean	SE	Mean	SE	Mean	SE
Retention of Correct Responses	.770	.026	.698	.020	.397	.028
Retention of Errors	.481	.034	.767	.019	.837	.023
Error Correction	.518	.034	.233	.019	.163	.023
New Errors	.230	.026	.302	.020	.603	.028

Table 3.
Experiment 2 Mean Proportion of Retention of Responses and Changing Responses as a Function of Feedback Type and Awareness.

	Aware					
	Feedback					
	True Feedback		No Feedback		False Feedback	
Proportion	Mean	SE	Mean	SE	Mean	SE
Retention of Correct Responses	.914	.075	.840	.056	.500	.080
Retention of Errors	.495	.099	.728	.054	.833	.065
Error Correction	.505	.099	.272	.054	.167	.065
New Errors	.086	.075	.160	.056	.500	.080
	Unaware					
	True Feedback		No Feedback		False Feedback	
Proportion	Mean	SE	Mean	SE	Mean	SE
Retention of Correct Responses	.750	.028	.678	.021	.38	.029
Retention of Errors	.480	.037	.773	.020	.839	.024
Error Correction	.520	.037	.227	.020	.164	.024
New Errors	.250	.028	.322	.021	.616	.030

Table 4.
Experiment 3 Mean Proportion of Retention of Responses and Changing Responses

	Preschool Participants			
	Feedback		No Feedback	
Proportion	Mean	SE	Mean	SE
Retention of Correct Responses	.970	.029	.871	.041
Retention of Errors	.088	.034	.807	.035
Error Correction	.933	.047	.135	.025
New Errors	.037	.023	.137	.029

	College Participants			
	Feedback		No Feedback	
Proportion	Mean	SE	Mean	SE
Retention of Correct Responses	.961	.013	.942	.016
Retention of Errors	.187	.048	.702	.016
Error Correction	.802	.048	.300	.054
New Errors	.038	.014	.060	.015

APPENDIX B. FIGURES

Figure 1.
Experiment 1 mean percent correct as a function of receiving feedback and test type.

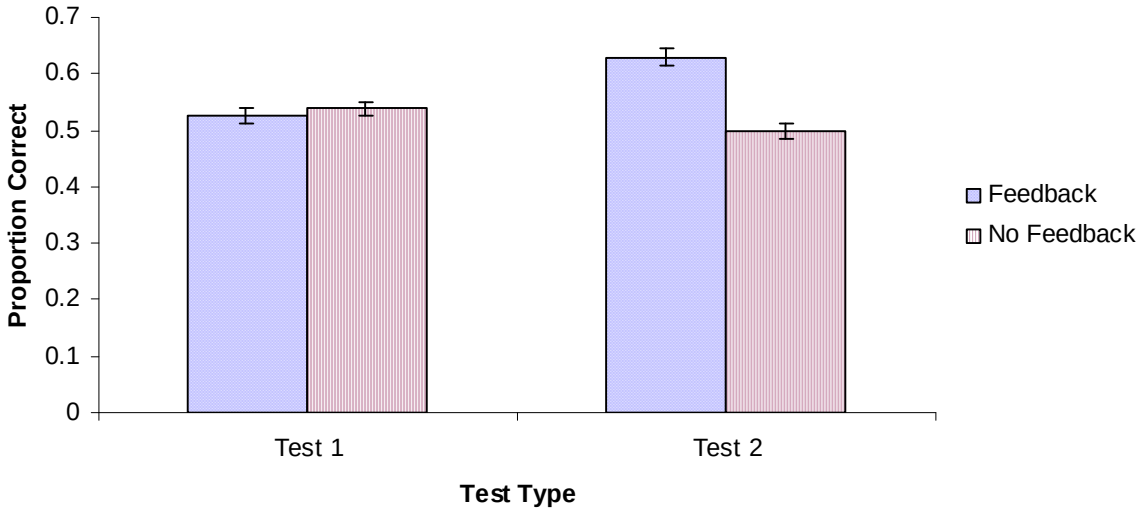


Figure 2.
Experiment 1 mean confidence rating as a function of receiving feedback and test type.

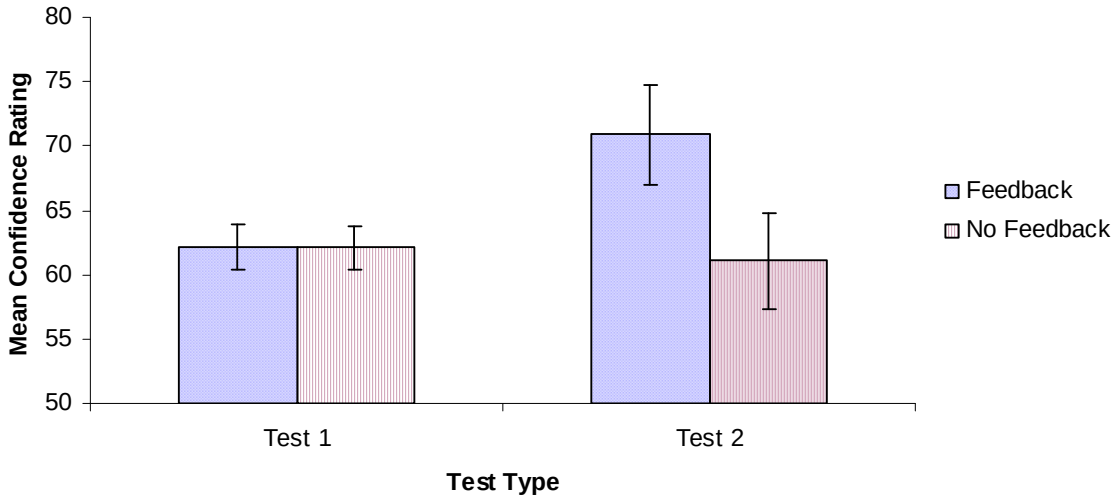


Figure 3.
Experiment 1 mean correct retention as a function of receiving feedback and confidence rating.

Figure 4.

Experiment 1 mean error correction as a function of receiving feedback and confidence rating.

Figure 5.
Experiment 2 mean proportion of following the feedback as a function of feedback type.

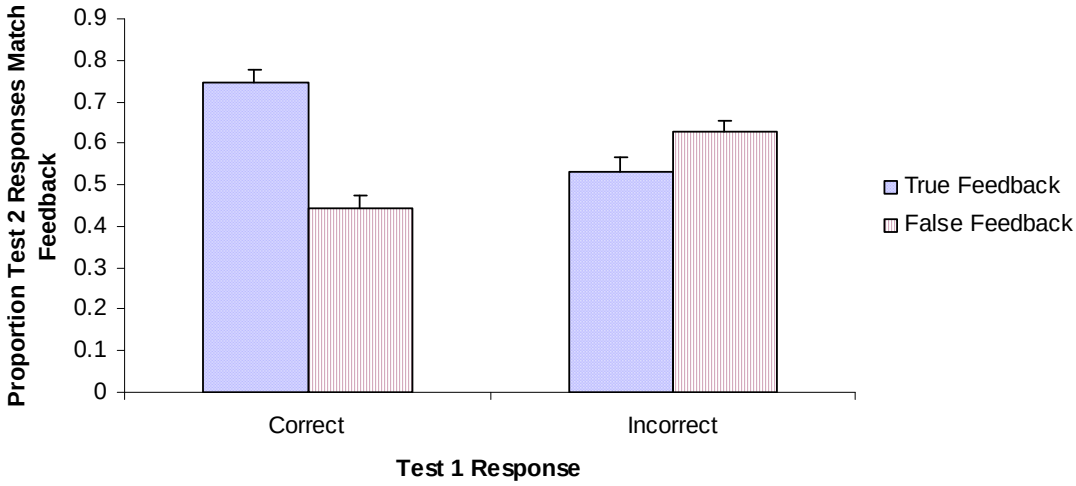


Figure 6.
Experiment 2 mean confidence rating as a function of feedback type and test type.

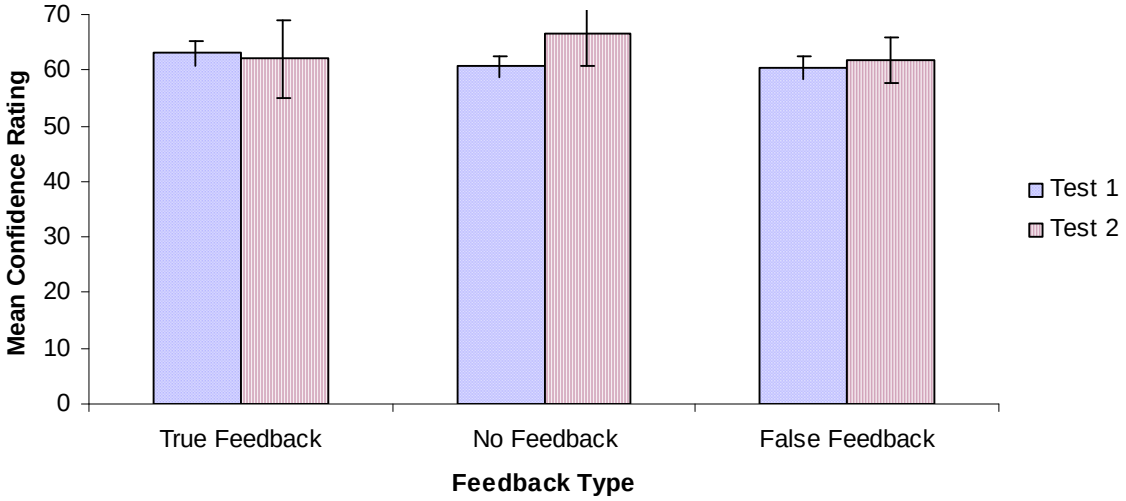


Figure 7.
Experiment 2 mean correct retention for high confident test 1 correct responses and for low confident test 1 correct responses.

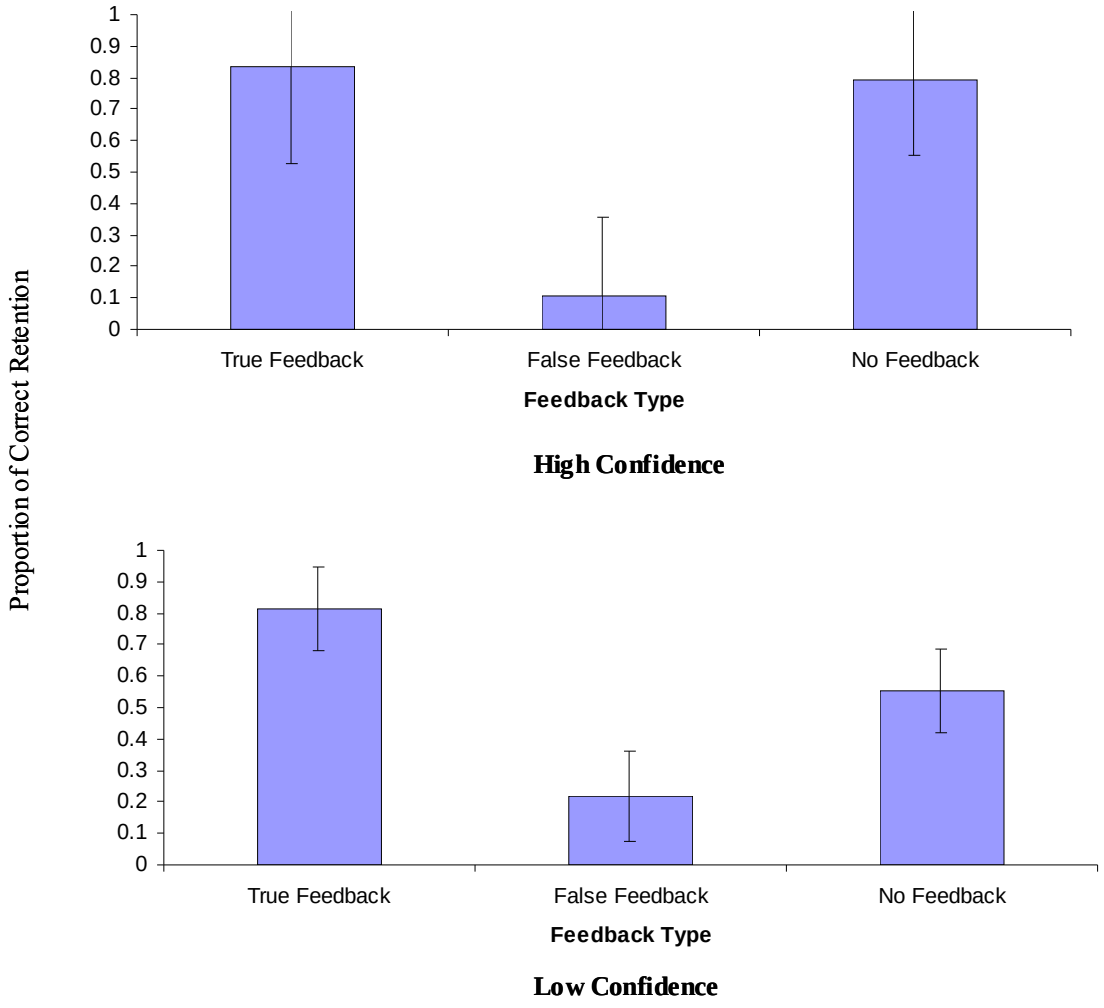


Figure 8.
Experiment 2 mean error correction for high confident test 1 errors and for low confident test 1 errors.

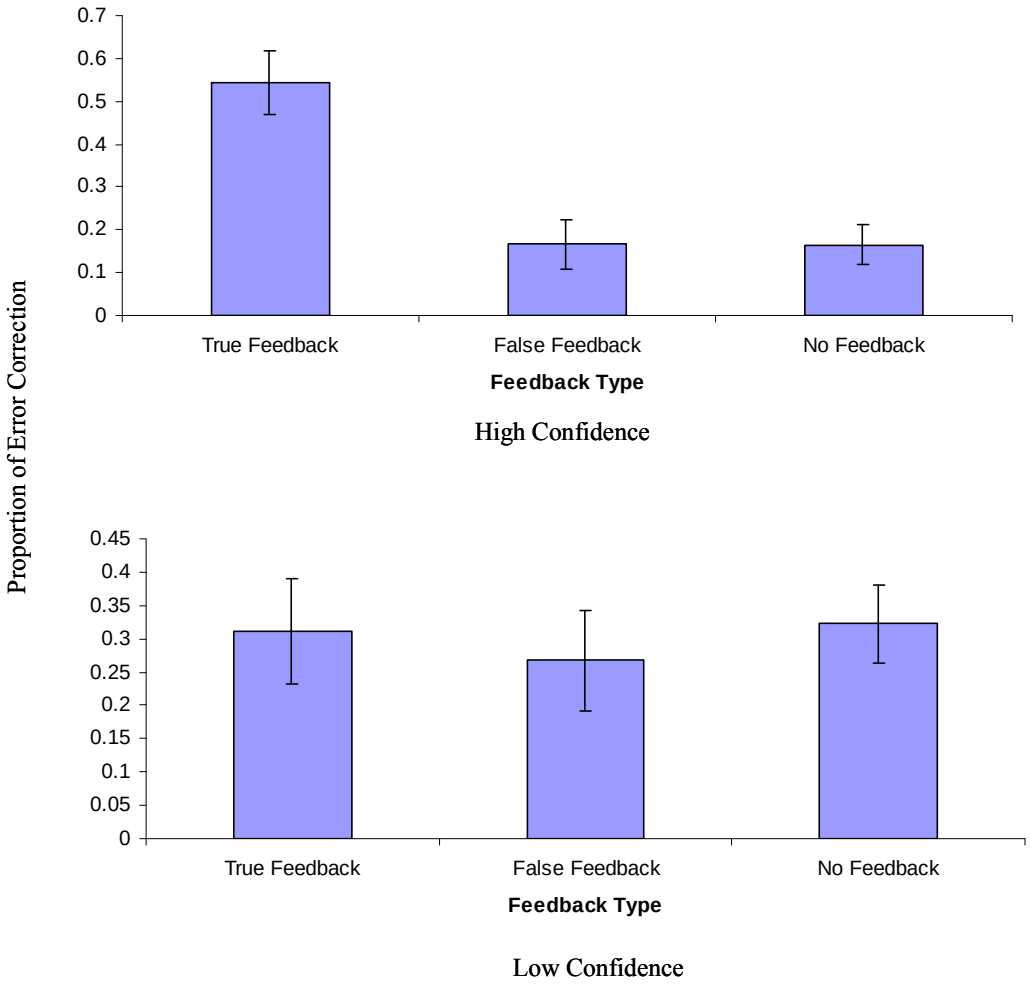


Figure 9.
Experiment 2 mean proportion correct as a function of feedback, test type, and awareness.

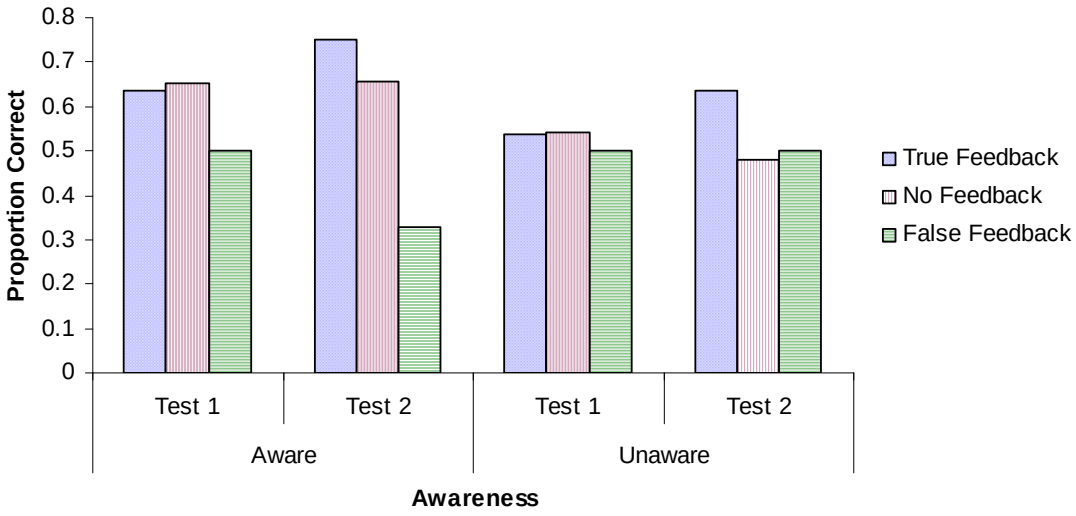


Figure 10.
Experiment 2 mean proportion of responses that follow the feedback as a function of receiving feedback, awareness, and Test 1 response.

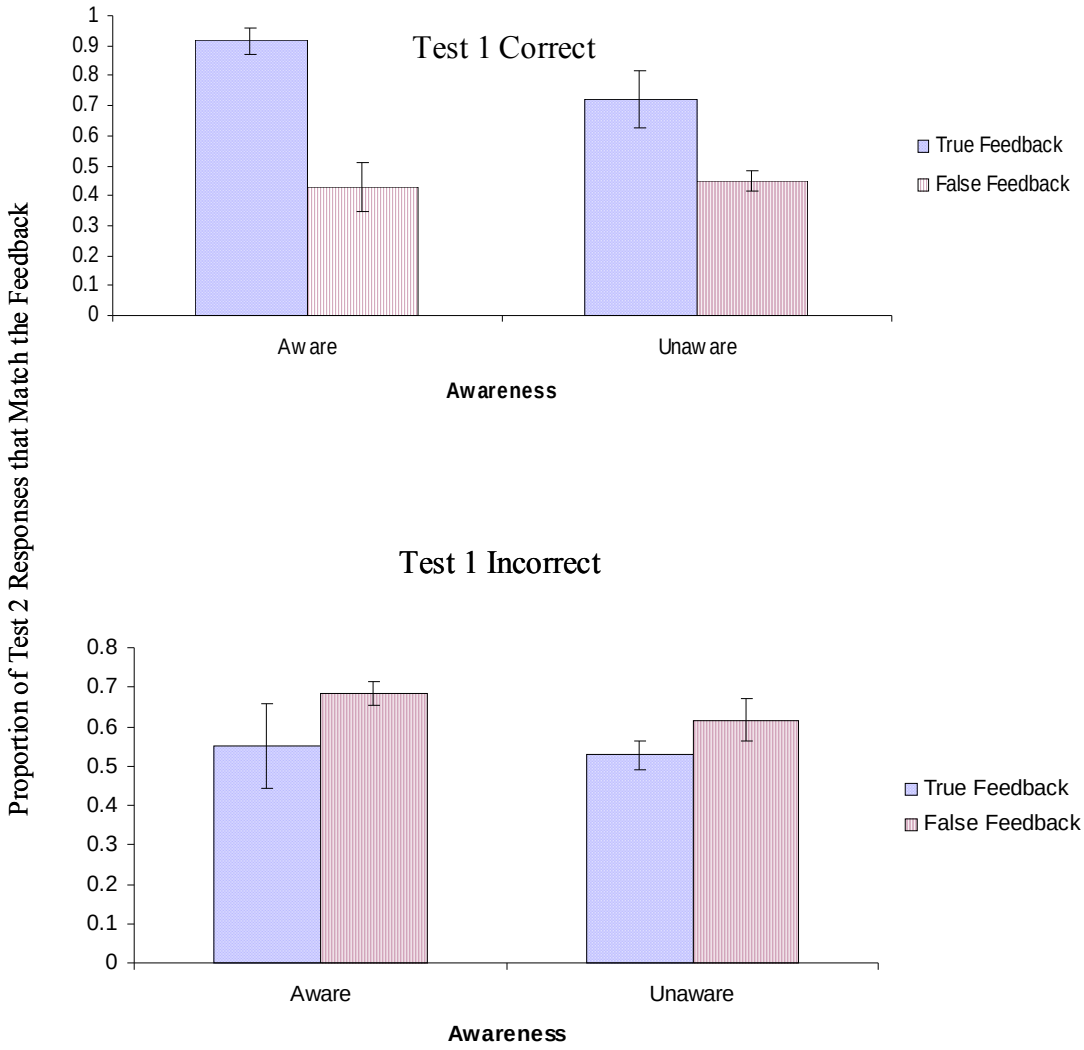


Figure 11.
Experiment 3 mean percent correct as a function of receiving feedback and test number for preschool participants and college participants.

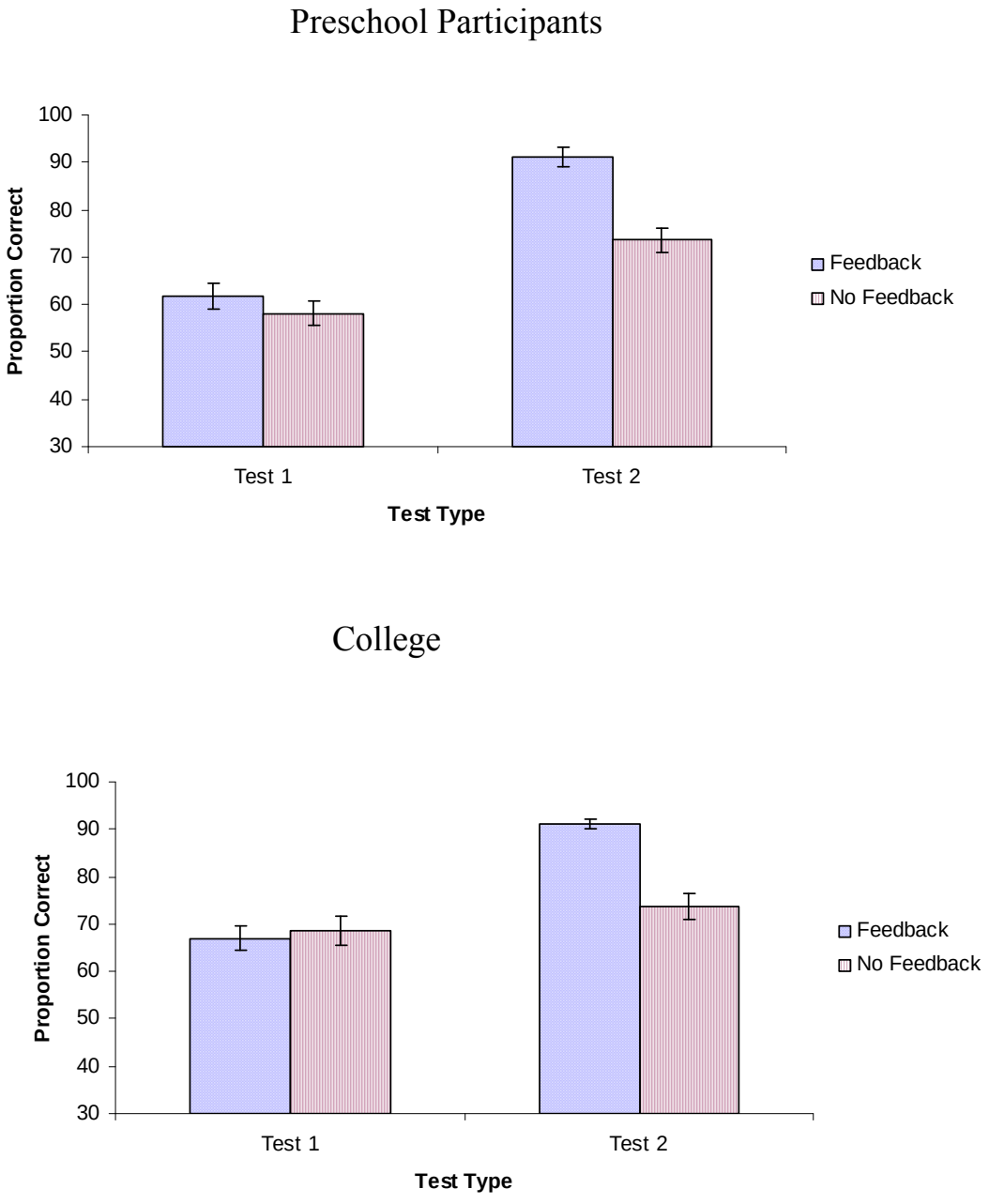


Figure 12.
Experiment 3 mean confidence rating as a function of receiving feedback and test type for preschool participants and college participants.

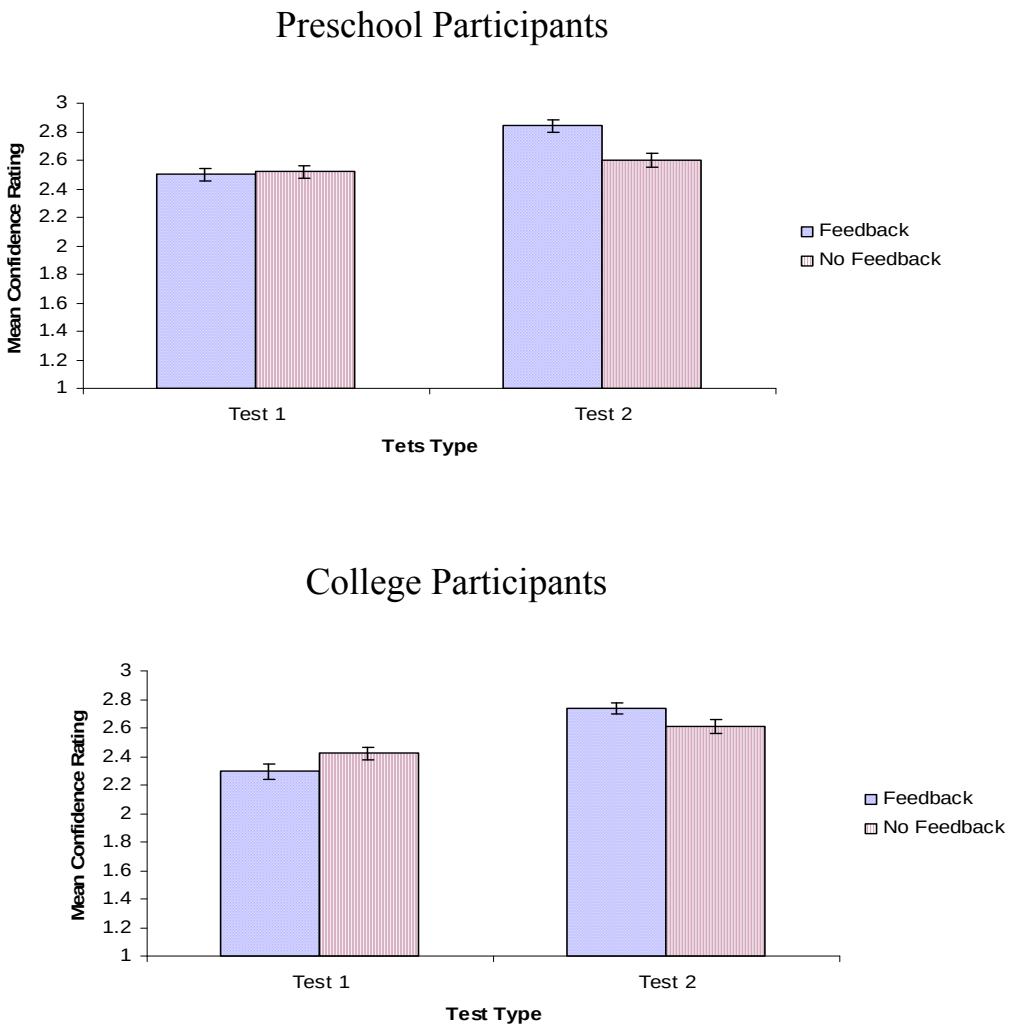


Figure 13.
Experiment 3 retention of correct responses for high and low confident correct responses as a function of receiving feedback for preschool participants.

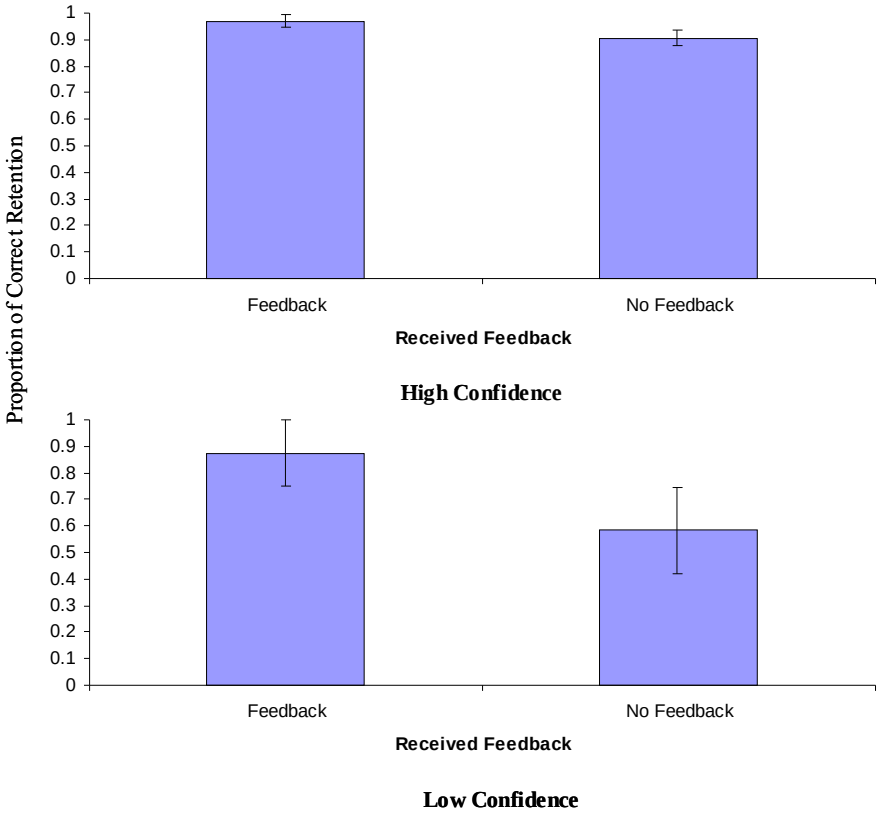


Figure 14.
Experiment 3 mean correct retention for high confident test 1 correct responses and for low confident test 1 correct responses for college participants.

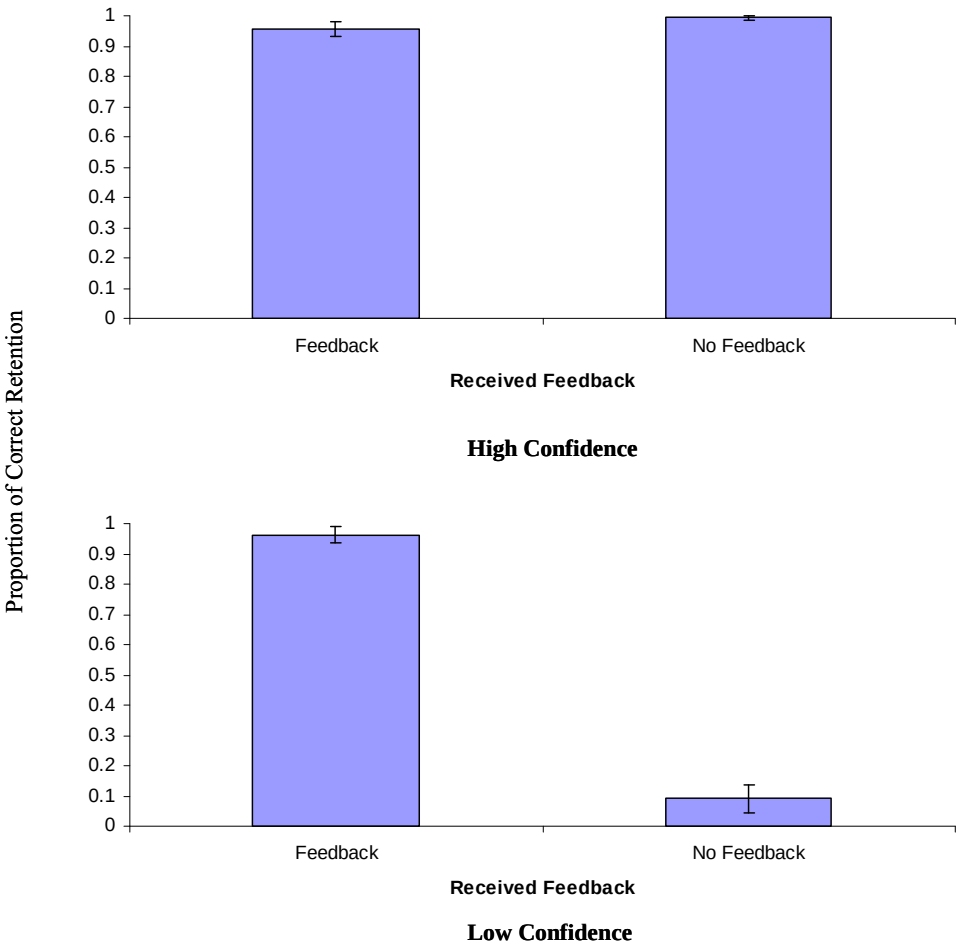


Figure 15.
Experiment 3 error correction for high and low confident errors as a function of receiving feedback for preschool participants.

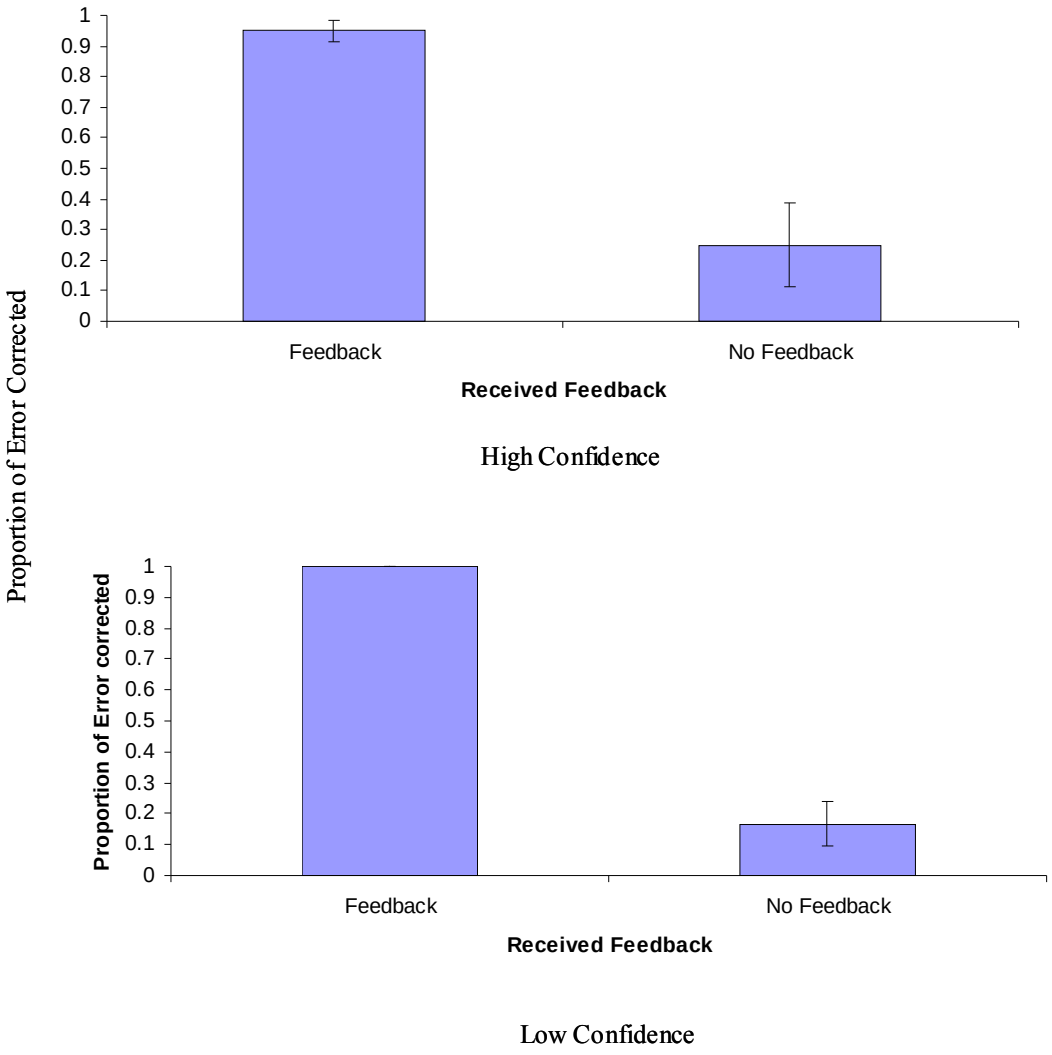
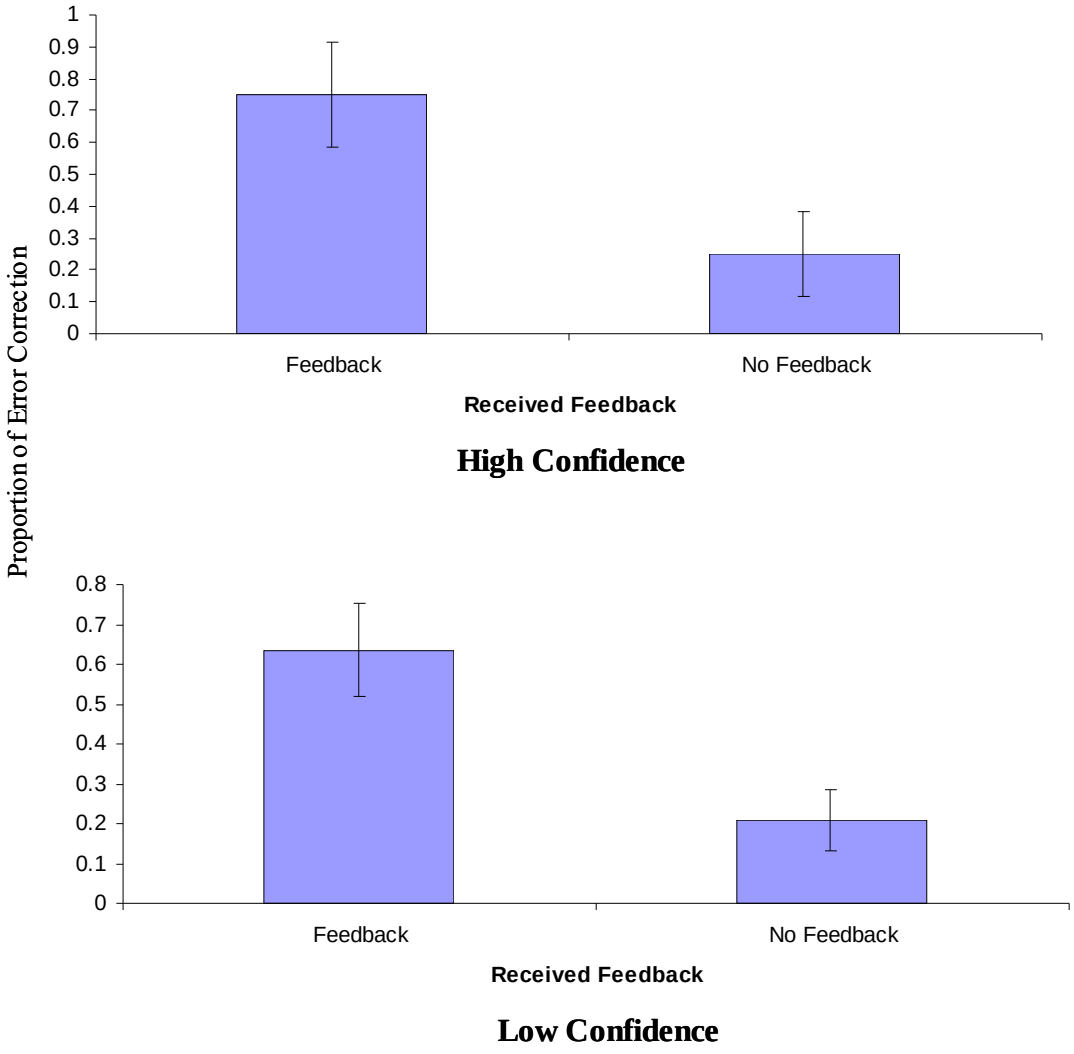


Figure 16.
Experiment 3 mean error correction for high confident test 1 errors and for low confident test 1 errors for college participants.



APPENDIX C. MATERIALS

Material 1.
List of Current Direction in Psychological Science Articles and Children's Books.

Current Direction Articles

Anastasio, P.A., Rose, K.C., & Chapman, J. (1999). Can the media create public opinion? A social identity approach. *Current Directions in Psychological Science, 8*, 152-155.

Eagly, A.H., Kulesa, P., Chen, S. (2001). Do attitudes affect Memory? Tests of the congeniality hypothesis. *Current Directions in Psychological Science. 10*, 5-9.

Garry, M., & Polaschek, D.L.L. (2000). Imagination and memory. *Current Directions in Psychological Science, 9*, 6-10.

Treiman, R. (2000). The foundations of literacy. *Current Directions in Psychological Science, 9*, 89-92.

Children Books

Hoff, S. (1958). *Danny and the Dinosaur*. New York: Harper Collins.

Kirwan, W. (2007). *Nobody Notices Minerva.* New York: Sterling.

Material 2.
Questions for the Current Direction in Psychology Articles.

(1) What is one of the most blatant examples of how the media can induce public opinion?

A. biased news coverage
B. "live" telecast of events
C. advertisements
D. selective censorship of news stories

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(2) What difference did Archer et al. (1983) find in the way men and women are typically portrayed in news photographs?

A. Men are often pictured in job-related roles, whereas women feature more prominently in home-related roles.
B. Photographs of men tend to be more close-up compared to that of women.
C. The facial expressions of men in photographs tend to be more solemn than that of women.
D. Men tend to be photographed alone, whereas photos of women tend to feature them in a group.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(3) Persons depicted in photographs high in "face-ism" tend to be rated as more _____ .

A. friendly
B. confident
C. trustworthy
D. Intelligent

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(4) According to Mullen et al. (1986), how was newscaster Peter Jennings different when discussing Ronald Reagan's 1984 campaign compared to when he discussed the campaign of Regan's political opponent?

A. Peter Jennings smiled more when discussing Reagan.
B. Peter Jennings used more hand gestures when discussing Reagan.
C. Peter Jennings used more hand gestures when discussing Reagan.
D. Peter Jennings looked directly at the camera more often when discussing Reagan.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(5) According to Gilens (1996), how can the media's portrayal of America's poor affect public perception of poverty?

A. The media's overrepresentation of African Americans in poverty can create the perception of more blacks in poverty than there actually are.
B. The media's portrayal of poor people as lacking in motivation can lead to less public support for social welfare and public assistance.
C. The media's portrayal of people in poverty as being lazy can increase negative attitudes.
D. The media's underrepresentation of certain groups in their portrayal of poverty can lead to those groups being neglected in social welfare policies.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(6) What is the hostile media bias?

A. The subtle effects of media portrayal on people's perceptions and opinion.
B. The media's influence on hostile and aggressive behavior.
C. The media's reinforcement of negative stereotypes of out-group.
D. People on both sides of a controversy perceiving the media as hostile to their group.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(7) Advertising that uses an attractive person to promote a product is relying on the _____ route of persuasion.

A. central
B. secondary
C. peripheral
D. fundamental

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(8) A study conducted by the authors (which involved subjects judging guilt/innocence of a fraternity member on charges of vandalism) found that the subject's tendency to side with one's in-group disappeared.

A. the subject was exposed to the opinion of an authority figure
B. the subject was exposed to evenly mixed opinion
C. opinions of others were homogeneous and perfectly correlated with group membership
D. the subject was given time to consider all evidence

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(9) Person's depicted in photographs low in "face-ism" also tend to be rated as

A. funnier
B. more dominant
C. whiter
D. heavier

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(10) Zuckerman & Kieffer (1994), Mullen et al. (1986), and Gilens (1996) all assumed the perceiver is an unbiased information processor who

A. passively accepts the views offered by the media.
B. assumed the perceiver is biased and motivated to filter the information according to
C. support the advertisers bias
D. deny the hostile media bias

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(11) The in-group influences the thinking of members of that group on a particular topic

A. when they have similar beliefs on that topic
B. even when they disagree on that topic
C. When the out-group is very outspoken against the in groups belief
D. When they have similar beliefs on all topics

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(12) More people who watched what network news station voted for Reagan?

A. Fox
B. CBS
C. NBC
D. ABC

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

Eagly, Kulsea, Chen (2001)

(1) What is the congeniality hypothesis?

A. People are motivated to avoid information that challenges their attitudes.
B. People's memories are biased in favor of information that agrees with their attitudes.
C. People selectively pay attention only to attitudinally agreeable information.
D. People tend to more elaborately process information that is inconsistent with their attitudes.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(2) The author's meta-analysis of research on memory for attitude-relevant information, what was the trend in results for later compared to early findings?

A. Early experiments showed that congenial information was less memorable than uncongenial information, whereas later research tended to yield a larger effect of congeniality on memory than early studies.
B. The results tended to be inconsistent regardless of whether the studies were early or more recent.
C. Participants in earlier studies tended to have weaker attitudes.
D. Later studies examined more variables than earlier studies.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(3) What is the likely cause of the different trend in findings for earlier vs. later studies?

A. Improvements in the procedures used to assess memory.
B. Participants in later studies tended to have less polarized attitudes.
C. Participants in earlier studies tended to have weaker attitudes.
D. Later studies examined more variables than earlier studies

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(4) What is the design procedure for a typical experiment looking at the congeniality effect?

A. Participants attitudes toward an issue are measured before and after presentation of information relevant to the issue.

B. Participants are presented with information that disagrees with their attitudes and their subsequent memory for that information assessed.

C. Participants with opposing attitudes toward an issue are presented with information on one or both sides of the issue, and their subsequent memory for that information assessed.

D. Participants are presented with information that agrees with their attitudes, and their subsequent memory for that information assessed.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(5) What did the authors propose could account for the weakness of the congeniality effect shown in experiments that were methodologically more rigorous?

A. Attitudes have little impact on memory.

B. people avoid information that challenges their attitudes.

C. People may mount an active defense and hence thoroughly process counterattitudinal information.

D. Participants had insufficiently strong attitudes.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(6). In their recent experiment (Eagly et al., 2000), the authors found what difference between congenial and uncongenial information?

A. Congenial information was recalled better than uncongenial information.

B. Participants had more prior knowledge of congenial information than uncongenial information.

C. Uncongenial information was better recalled soon after the message was presented, whereas congenial information was better recalled after a delay.

D. Uncongenial information elicited more thought and attention than congenial information.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(7) The authors propose that to persuade people to accept a position that is highly divergent from their own attitudes, it might be best to

A. use an incremental approach whereby each exposure to uncongenial information produces only a small amount of change.
B. expose them to large amounts of uncongenial information at one go.
C. employ an authority figure to promote the counterattitudinal position.
D. encourage them to think global thoughts concerning the issue rather than differentiated thoughts.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(8) According to dual-process theories of social judgment, a recipient who is lacking in motivation and capacity will likely adopt what type of approach when f aced with uncongenial information?

A. yield and capitulate to the counterattitudinal viewpoint
B. adopt an active resistance and confront the uncongenial information
C. adopt a passive, avoidant approach and process the information less
D. react emotionally and dismiss the information outright

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(9) What is the usual rationale for predicting better memory for information that is attitudinally congenial than for information that is not congenial?

A. people are motivated to defend their attitudes against material that challenges them
B. people do not understand the opposing argument as well; therefore, it is more difficult to remember opposing information
C. People are more familiar with congenial information; therefore, it is easier to remember congenial information
D. People remember hostile information better because it is surprising

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(10) The congenial effect is

A. is well supported in the literature
B. received better support in later studies
C. received better support in early studies
D. is the one of the strongest effects in psychology research

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(11) The authors' current experiments studied the congenial effect using 3 social issues (abortion, gays in the military, and the death penalty). They found the congenial effect

A. stronger with the abortion issue
B. non-existent in all 3 social issues
C. stronger with the gays in the military issue
D. stronger with the death penalty issue

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(12) Compared to recipients of a uncongenial message, when asked to list the thoughts that had come to mind as they listened to the message recipients of a congenial message

A. generated more thoughts that were relevant to the message
B. generated less thoughts that were relevant to the message
C. generated equivalent thoughts that were relevant to the message
D. generated no thoughts that were relevant to the message

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

Garry and Polaschek (2000)

(1) Loftus (1993) was the first systematic study to show what?

A. detailed false memories for a whole event could be implanted
B. emotional events tend to be particularly salient and memorable
C. counterfactual thoughts can affect people's judgment of outcomes
D. people tend to misremember childhood events

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(2) According to Sarbin (1998), what strategy do people rely on when they try to remember an event that they do not remember?

A. fabrication of the details
B. exhaustive search of their memory store
C. imagination of the event
D. look for retrieval cues in the environment

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(3) Subjects become more confident they have experienced a counterfactual event after they imagine the event. This is called

A. false memory
B. imagination inflation
C. illusory vividness
D. confirmatory bias

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(4) What is source confusion?

A. Confusing details from an imagined event with details from an experienced event.
B. Forgetting what the source of a memory was.
C. When the vividness of a memory is no longer a good indicator of its veracity.
D. Misattributing content of a memory to the wrong source.

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(5) Two mechanisms have been proposed to account for the boost in confidence of having experienced an imagined counterfactual event. One is source conclusion, the other is _____ .

A. strength of memory trace.
B. recollection.
C. vividness of memory.
D. familiarity

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(6) According to Heaps and Nash (1999), which of the following factors predicts people's tendency to become more confident that they have actually experienced an event after imagining the event?

A. Their susceptibility to influence of an authoritative person
B. The vividness of their mental imagery
C. Their predisposition to hypnotic suggestion
D. Their arousal to emotional stimuli

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(7) One might be tempted to regard the confidence boosting effect of imagining an event as merely the statistical phenomenon of

A. regression towards the mean
B. restriction of range
C. homogeneity of regression
D. a spurious correlation

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(8) Why do the findings of memory-related effects of repeatedly imagination have clinical implications?

A. Because patients might be imagining their disorder/illness
B. Because various psychotherapy techniques involve imagining situation and actions
C. Because the therapist may find it difficult to distinguish reality from imagination
D. Because repeated imagination of events can lead to hallucinations

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(9) Counter-factual thinking is

A. Thinking "What if" or "If only I had done something else"
B. Believing in a false memory
C. Using fictional information to make a decision
D. List all possible arguments and their counter-arguments before making a decision

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(10) How many subjects developed a false shopping-mall memory.

A. 0%
B. 100%
C. 25%
D. 50%

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(11) Loftus (1993) and Hymann and Pentland (1996) showed what?

A. Counterfactual thinking leads to very accurate memories
B. Thinking about an event is not as important as imagining an event
C. False memories rarely occur
D. false memories can be created when people think about childhood events in an attempt to remember them

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(12) What is the most arguable of the individual differences considered in the false memory literature?

A. Susceptibility to hypnosis
B. Gender
C. Age
D. Susceptibility to suggestion

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

Treiman (2000)

(1) What is the alphabetic principle?

A. appreciating that many languages, in the written form, use a set of symbols or letters
B. appreciating how the letters in printed words
C. appreciating that there are some rules in how letters can be combined in the spelling of words
D. appreciating how the spelling of words can be inconsistent

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(2) What is the phoneme?

A. basic sound unit of a language
B. the sound structure of a language
C. a syllable
D. a cluster of consonants

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(3) A syllable can be subdivided into

A. consonant clusters
B. vowel clusters
C. letter segments
D. onset and rime

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(4) Studies have shown that training in _____ can improve reading and spelling ability in children.

A. the names of letters
B. analyzing linguistic structure
C. phonological awareness
D. how to spell their names

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(5) What has been the implicit assumption about how children learn letter names and letter sounds?

A. they learn them via imitating adult speech
B. they learn them unconsciously when listening to adults speak
C. they learn them via experimentation with different sounds
D. they learn them via rote memorization

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(6) Recent studies by Treiman et al. have found that an important determinant of knowledge of letter-sounds is

A. whether the letter's sound occurs in the name of the letter
B. whether the letter is voiced or unvoiced
C. the place of articulation of the sound
D. the spelling of the child's name

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

7. Young Joe is more likely to now the _____ of the letter 'j' than Alice or Tom.

A. place of articulation
B. phoneme
C. name
D. sound

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

8. There is a widespread view that young children are purely _____ readers, memorizing associations between whole printed words and their spoken form.

A. logographic
B. autonomic
D. articulate
E. biographic

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(9) A child who possess phonological awareness knows that "bat" and "got" are alike

A. they both have a vowel in the middle of the word creating a similar phoneme
B. the both have 3 letters creating a similar phoneme
C. because they share the 3rd phoneme
D. they are both 1 syllable

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(10) Letter knowledge includes ALL of the following EXCEPT

A. knowledge of letter names
B. ability to retrieve this information quickly
C. knowledge of letter sounds
D. knowledge of letter stems

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(11) Recent studies

A. support the idea that children learn the sounds of letters through rote memorization
B. show letter knowledge and phoneme awareness are related
C. found letter-sound knowledge is largely determined by if the sound of the letter is in its name
D. support the idea that poor readers have good letter knowledge but poor phoneme awareness

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

(12) What does logographic mean?

A. readers memorize associations between printed words and spoken words in a rote fashion
B. Readers associate sounds with letters
C. Readers associate letter names with their sounds
D. Readers memorize the look of the letter

Not at all Sure		Somewhat Sure		Strongly Sure
1	2	3	4	5

Material 3.
Example of Experiment 1 Feedback.

Loftus (1993) was the first systematic study to show what?

A. Detailed false memories for a whole event could be implanted.
B. Emotional events tend to be particularly salient and memorable.
C. Counterfactual thoughts can affect people's judgment of outcomes.
D. People tend to misremember childhood events.

Your answer was: (D) People tend to misremember childhood events

The CORRECT answer was: (A) Detailed false memories for a whole event could be implanted.

Material 4.
Example of Experiment 2 Feedback.

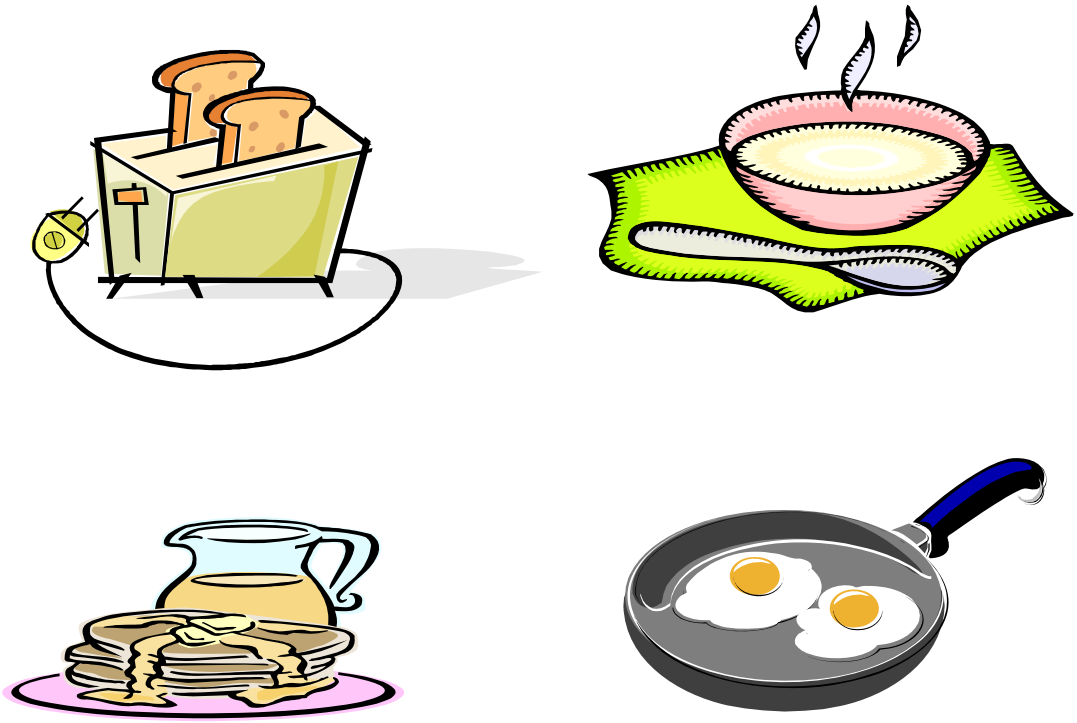
Loftus (1993) was the first systematic study to show what?

A. Detailed false memories for a whole event could be implanted.
B. Emotional events tend to be particularly salient and memorable.
C. Counterfactual thoughts can affect people's judgment of outcomes.
D. People tend to misremember childhood events.

Your answer was: (D)

The CORRECT answer was: (A)

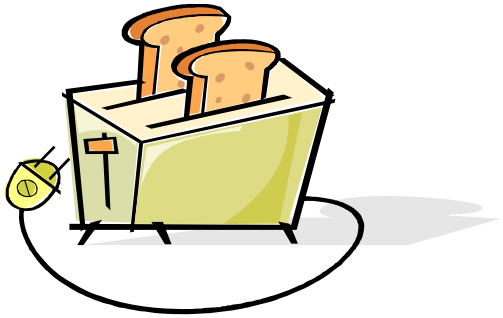
Material 5.
Example of Children's Material Questions Used with Preschool Participants.



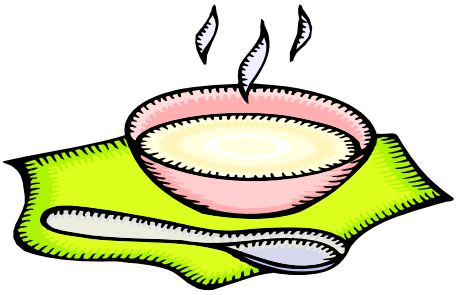
Material 6.
Smiley Confidence Materials.

Material 7.
Example of Children's Material Questions Used with College Participants.

What did Minerva's baby sister have for breakfast?

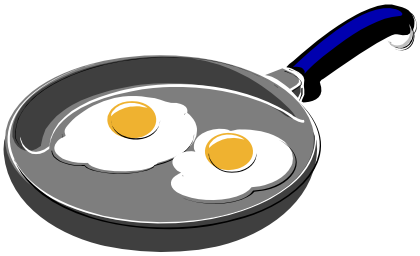


toast



oatmeal

pancakes



eggs