METACOGNITIVE ACCURACY FOR FACE-NAME PAIRS USING A DUAL-PROCESS APPROACH TO JUDGMENTS OF LEARNING

By

Taylor Michael Curley

A paper submitted in partial fulfillment of the requirements of the Honors College in the Department of Psychology

Examining Committee:	Approved by:
____________________________ Karen Daniels, Ph.D	____________________________ Jeffrey Toth, Ph.D Faculty Supervisor
____________________________ Jeanne Kemppainen, Ph.D, RN	____________________________ Richard Ogle, Ph.D Department Head
____________________________ D. Jeanne Swafford, Ph.D Honors Council Representative	
____________________________ Katherine E. Bruce, Ph.D Director of the Honors College	

University of North Carolina Wilmington

Wilmington, North Carolina

April, 2012

TABLE OF CONTENTS

ABSTRACT

This project explores the nature of memory and metamemory for face-name pairs using a dual-process approach to Judgments of Learning (JOLs) initially described by Daniels, Toth, and Hertzog (2009). Young adults (n = 48, mean age = 19.15) studied a set of 96 face-name pairs. Interleaved with these study trials were JOL trials in which participants were asked to rate the likelihood that they would remember a name when later prompted with its corresponding face and the first letter of the name. This rating occurred either immediately after studying a face-name pair (immediate JOL) or after a delay of six or seven other face-name pairs (delayed JOL). Then, at test, participants were shown the faces along with the first letter of the corresponding names and asked to recall the names that were paired with the faces. When a name was produced, participants were further asked to describe whether they recollected the name as being paired with the face ("Recollect"), whether the name came to mind automatically without any recollective details ("Know"), or whether they had no memory for the face-name pairing ("No Memory"). Results showed a small but significant difference in memory accuracy such that name recall was better in the delayed JOL condition, compared to the immediate JOL condition. Results also showed that JOL accuracy was significantly higher for face-name pairs in the delayed compared to immediate JOL condition, thus extending the delayed JOL effect to the recall of names cued with faces. Finally, replicating Daniels et al. (2009), name recall associated with Recollection was found to provide the most reliable diagnostic of future memory performance, relative to Know and No Memory judgments, although this pattern varied in the immediate and delayed JOL conditions. Overall, the data suggest that JOL accuracy for face-name pairs is primarily driven by recollective-based memory and that the delayed-JOL effect can be replicated with face-name pairs.

ACKNOWLEDGMENTS

I would like to acknowledge all of the graduate and undergraduate students of the Aging and Cognitive Training (ACT) Lab, past and present, for their encouragement and assistance to this project and others during my time at the university. Without their help and guidance, I would still be lost to the endless amounts of papers written on cognition and my project would have never come to fruition. Of these, I would especially like to thank Caitlin and Matthew for assisting me with the details of the project and giving me some idea of how to use E-Prime.

I would also like to thank the members of my examining committee, Dr. Daniels and Dr. Kemppainen, as well as the staff of the Honors College for their guidance during the synthesis of this project, including Dr. Bruce and Dr. Swafford. The questions and ideas produced by these facilitators have proven to be incredibly enhancing to my research as well as my own way of thinking.

An acknowledgement would not be complete, of course, without the inclusion of my family and friends, all of whom have been there to support me ever since I decided to pursue an Honors Thesis. My special thanks goes out to Jessi, Mom, Dad, Alex, Caitlin, and Philip who have always been there to provide the love and encouragement that I needed.

Finally, I would like to acknowledge my thesis advisor and mentor, Dr. Jeffrey Toth, for continuing to encourage and guide my growth as a cognitive scientist. When I first walked into his office to inquire about working in the ACT Lab, I was completely oblivious to the amount of knowledge and work that would be required of me, but as I have come to the end of my undergraduate education, I can see that the amounts of hard work I encountered were completely essential to my growth as a student and a cognitive scientist. Dr. Toth has always encouraged me to be the best researcher I possibly could, and for this I thank him tremendously.

LIST OF TABLES

LIST OF FIGURES

METACOGNITIVE ACCURACY FOR NAME-FACE PAIRS USING A DUAL-PROCESS APPROACH TO DELAYED JUDGMENTS OF LEARNING

INTRODUCTION

One of the most significant cognitive processes that we use regularly is our ability to remember names and faces together. Whether the situation that demands the task is social (such as a party) or just for entertainment (like remembering an actor in a movie), the act reflects cognitive processes that are frequently used, but never fully examined by the observer. While the importance of memory for peoples' names and faces is immediately recognizable, the extent to which we use these processes and the ways that we can improve them continue to be explored by researchers of cognition, memory, and psychology in general.

The purpose of this research is to continue the inquiry into the underlying cognitive forces that drive our memory for names and faces and the methods that we can use to increase our memory for them. Much of the past research into memory for these items has garnered many interesting results and suggests that learning pairs of faces and names is considerably more difficult than just names, faces, or bits of personal information by themselves (McWeeny, Young, Hay, & Ellis, 1987; Watier & Collin, 2011). The difficulty in this research, of course, lies in the fact that names and faces are processed by two different parts of the brain, as indicated by ERP research and the neural correlates associated with each process (MacKenzie & Donaldson, 2009; Weise, Schweinberger, 2008; Pickering & Schweinberger, 2003); but, despite this, the two processes remain intricately connected and help govern our social cognitive functioning on a daily basis.

By examining metacognition as it relates to name-face pairs, this project will explore two of the most important skills that an individual uses when remembering a person's name and face: predictive and retrospective memory. If individuals are prompted with the task of remembering a name of importance, will they be able to accurately predict how well they will later remember that name? Or will they incorrectly predict their ability to remember and remain unable to recall the person's name when they see them next? These are just a few of the questions underlying the theoretical objectives of this research.

In the current experiment, two main focuses of inquiry concerning memory for names and faces were examined. The first is to examine the functions governing predictive memory accuracy in memory for names and faces while the second is to examine one particular method that may improve accuracy for these items. To examine metamemory for name-face pairs, we employed the dual-process judgment of learning (JOL) technique used by Daniels, Toth, and Hertzog (2009; see also Toth, Daniels, & Solinger, 2011) wherein JOL ratings were analyzed in terms of whether participants retrieved the name in a consciously controlled fashion ("Recollection" or R), or whether the name was retrieved more automatically, unaccompanied by memory for details of prior encounters with the name-face pair ("Know" or K). We expected to find that name recall accompanied by recollective detail would be the primary source of JOL accuracy for the face-name pairs, whereas name recall unaccompanied by recollective detail ("Know" or "No Memory (N)" responses) would contribute little to JOL accuracy.

To examine the methods that increase our memory accuracy for names and faces, JOLs for half of the studied stimuli were immediate, meaning that the slide requesting a JOL occurred immediately after the slide initially presenting the face-name pair. The other half of the stimuli were "delayed," meaning that the second judgment slide was separated by a given amount of time. In this, we hoped to find both increased memory for the delayed items, as well as a delayed-JOL effect (Nelson & Dunlosky, 1991), which shows increased JOL accuracy for delayed, as compared to immediate, items. We calculated this effect using gamma correlations as well as a "backsorting" procedure whereby JOL ratings from the study phase were analyzed in terms of the R/K/N judgment made at test (Daniels et al., 2009). Finding this pattern would indicate that metamemory performance is increased when items being remembered are shown at an initial memory task and again at a relatively later time. Overall, this will show the importance of recollection rather than more automatic forms of memory in the monitoring of name-face pairs, and will also examine a potentially useful strategy to increase memory performance with face-name pairs. These two results would both replicate the findings of Daniels et al. as well as present new research to the field of metamemory for names and faces.

COGNITION AND METAMEMORY

Historically, metacognition has always been acknowledged, though left unnamed for hundreds of years. Even early philosophers such as Descartes were amazed by the phenomenon and theorized on the process. According to Metcalfe (2000), Descartes was enthralled with basic ideas underlying metacognition: “It was not so much thinking that was indisputable to Descartes, but rather thinking *about* thinking. He could not imagine that the person engaged in such self-reflective processing did not exist” (2000, p. 197). These ideas of self-reflection were not systematically examined again until the late nineteenth century when introspection and psychology were first introduced. Some researchers contend that early figures such as Borden Parker Bowne (1886) and Henry James (1890) attempted to study the monitoring of learning, but with limited success (Cavanaugh & Perlmutter, 1980). James was, however, successful in researching a memory phenomenon that would later become important to the development of metamemory as a distinct area of inquiry: the tip-of-the-tongue (TOT) phenomenon, or a memory that feels close to being recalled yet still remains elusive to the individual (James, 1950).

Metacognitive theory as is it recognized today only began to take shape in the mid-twentieth century. Hart (1965) began researching the nature of memory and self-monitoring for instances of recall in which the individual cannot produce the desired information, yet still feels strongly about knowing the answer and their ability to recognize the correct item. To measure these instances of memory monitoring, Hart introduced the idea of feeling-of-knowing (FOK), or a self-diagnostic of "what is stored in memory when the retrieval of a memory item is temporarily unsuccessful or interrupted" (1965, p. 214). He examined this state by testing participants on their general knowledge and then asking them to rate their confidence of how well they knew the answer beside the question. Though Hart's research introduced empirical aspects to what we now know as "metacognition," the phenomenon still had not taken a definite form.

Five years after the introduction of FOKs, Tulving and Madigan (1970), in their comprehensive review of memory theory, concluded their paper by stating that memory research needed to "start looking for ways of studying, and incorporating into theories and models of memory, one of the truly unique characteristics of human memory: its knowledge of its own knowledge" (p. 477). The authors did not explicitly name this phenomenon as "metamemory," nor was the term used much in the years following, but the article is still considered to have inspired much of the metacognitive research to come.

Levels of processing

Some of the first research to implicitly recognize metamemory theory in the 1970's was studies on levels of processing (LoP). Craik and Lockhart (1972) acknowledged in the LoP theory that the representation of incoming information from an individual's environment depends on the type of analysis they performed on it. They also theorized that the amount of information that will be retrieved is dependent on the nature of the "code," such that information encoded conceptually will be better remembered than information encoded perceptually (Craik & Lockhart, 1972; Jacoby & Craik, 1979).

In LoP studies, other researchers have acknowledged the role of self-monitoring in encoding and analysis. Ornstein and Corsale (1979) noted that incoming information is processed by a series of "analyzing structures" that serve as interpretive filters, showing that knowledge is implicit in the LoP paradigm. While "deep" processing is greatly affected by a person's preexisting knowledge structures, biases and mnemonic strategies also play an important role. LoP research has also found that participant knowledge about retrieval influences the depth of the processing and enhances memory at the encoding stage (Hunt & Rawson, 2011). Furthermore, recent studies have found that recollective distinctiveness (rather than familiarity) is the most indicative of deeper processing (but see Toth, 1996), though the underlying reason as to why deeper processing is more distinctive still remains debated in the academic literature (Gallo, Meadow, Foster, & Johnson, 2008).

Theoretical framework for metacognition

The most influential boost in research on metacognition came decades after the LoP paradigm emerged and was introduced in Nelson and Narens' (1990) widely influential paper on metacognition. In this, they introduced a dynamic theory of metamemory that separates the phenomenon into three principles (Figure 1). The first principle states that there are at least two

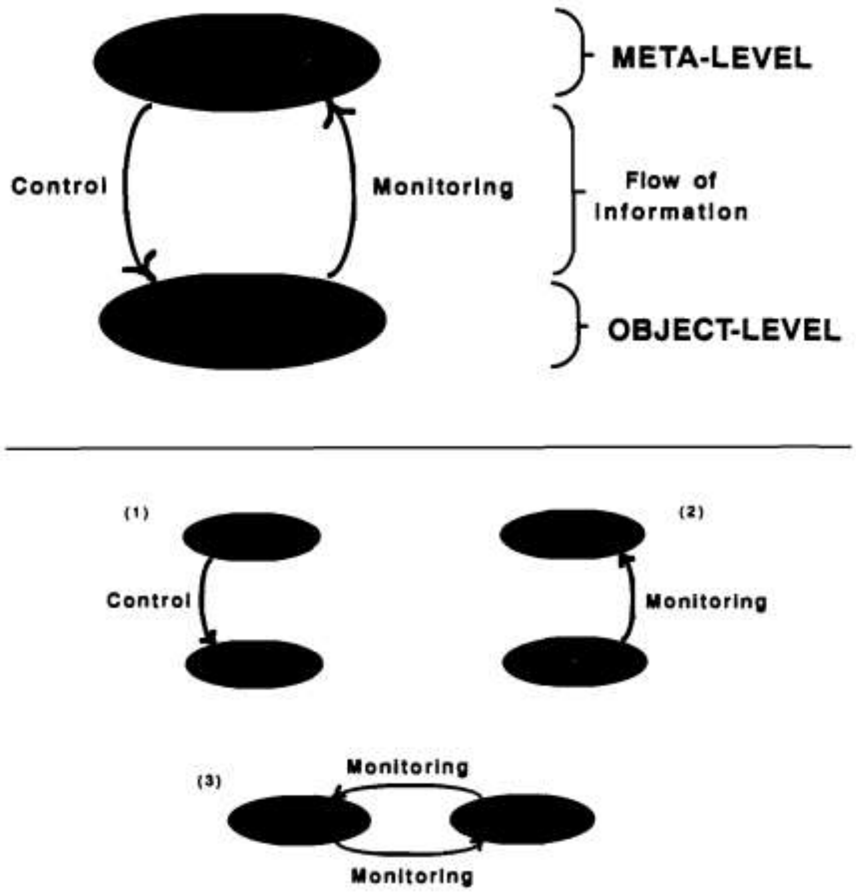

Figure 1. Illustrates the relationship between the object-level and the meta-level and the flow of information in the metacognitive framework (top half). The lower three figures represent theories of memory that disregard metacognition, accentuating the fallacies of the models that show control of information without feedback (1), knowledge of the information present but no control (2), and symmetrical modalities in which neither is meta-level to the other (3). Adapted from Nelson and Narens (1990).

different interrelated levels of metamemory: the object-level (what we would think of as memory and cognition) and the meta-level (metamemory and metacognition). The second principle states that a mental representation or abstract idea of the object-level exists in the meta-level, though not the reverse. This idea stems from previous research that suggests that monitoring systems must emulate the systems they wish to regulate, i.e. metamemory is an emulation of its controller, memory (Conant & Ashby, 1970). Lastly, the third principle postulates that "monitoring" and "control" are the "dominance relations" between the object- and meta-levels, depending on the flow of information (Nelson & Narens, 1990, p. 127). These researchers defined control as the mechanism that the meta-level uses to modify either the processes directing information to the object-level or the object-level itself. Monitoring is described as how the object-level communicates with the meta-level. Since the meta-level is a representation of the object-level, its state can change to reflect the information sent to it. Since the object-level does not model the meta-level, the object-level does not undergo any modal changes. The implications of the research, then, encourage further exploration into the idea that systematic introspection can change system behavior. The research presented here influenced the beginning of investigations into JOLs and help shape the current theories of metamemory that are still around today (Nelson, Dunlosky, Graf, & Narens, 1994).

MEASUREMENTS OF MEMORY MONITORING

Before the current methods of measuring metamemory (such as the ones used in this experiment) came to fruition, they underwent a long transformation of maturation and growth. The first empirical analysis of memory monitoring came with Brown and McNeill's (1966) investigation of the tip-of-the-tongue (TOT) phenomenon. Here, the TOT event was referred to as the state in which one feels that they know some target information and that it is on the verge of being accessed, but they cannot recall it. This would lead to eventual recall of the sought-after information, or the recognition that memory retrieval has failed (1966). To test this, the researchers compiled a list of words that are rarely used in conventional speech, but not so uncommon that the participants would not know what the word meant. During the experiment, the researchers would read aloud the definitions of these words to participants. If the individuals felt as if they were having a TOT experience, then they would fill out an information sheet describing their experience. To measure the reliability of the TOT state as a memory monitor, the researchers examined the correlation between the reported *believed* number of syllables and the actual number of syllables in the target word, which proved to be significant. They also looked at a number of other factors associated with memory for the word that could indicate the strength of the TOT experience and its accuracy in predicting the word the person was trying to remember. Overall, the researchers found that the TOT state is a very good monitor of episodic memory (1966). Despite the success of this initial predictive memory task, the authors acknowledged that there were some fundamental problems with the analysis. To start, only a few of the words in each trial actually evoked a TOT state, making the test a highly unreliable measure. The reliability of the participants' responses also presented a problem since the predictive accuracy for a word in a TOT experience cannot be examined consistently within subjects. Given these problems, Brown and McNeill's research was self-described as a "fragmentary data problem" (1966, p. 328).

Another memory phenomenon that was essential to the development of our current thinking about metamemory was first described by Hart (1965) and called the feeling-of-knowing (FOK), as described earlier in this paper. For their methods of measurement, the researchers asked the participants to rate how well they believed they knew the material. In the first experiment, this was done by completing a common knowledge test and answering “yes” or “no” to questions they could not answer as to whether they felt like they would know the correct answer if provided on a list. After the knowledge test, the subjects were then asked to go back through the test again, but this time with four answer choices listed. For this experiment, the data was analyzed through two measures: a common hit/miss analysis as to whether the question was answered correctly and another assessment of predictive memory accuracy. An answer that was left blank in the first task, given a “yes” judgment, and then correctly recollected was labeled as an FK hit. If the participant judged “no” for a blank item and subsequently failed to provide the correct answer, it was labeled as an $\underline{\text{FK}}$ hit (1965). In this first experiment, the researchers found that while individuals were mostly accurate judges of what they did know (FK), they were very poor judges of what they believed was *not* in their memory ($\underline{\text{FK}}$).

In the second experiment, the participants were asked to perform the same tasks as in the first experiment, but with a rating system that was modified to measure different degrees of uncertainty (Hart, 1965). In this, the levels of accuracy that the participants reported were graded from 1-6, 1 being the lowest level of certainty and 6 being the highest. This time, when the data was analyzed using the dichotomous procedure in Experiment 1, both types of “hits” (i.e. when the individual gave a higher rating for an item that was subsequently produced or a low judgment for one that was not produced) were both found to be reasonably high predictors of memory performance. Furthermore, the proportions of the correct answers as a function of each of the numbers on the scale were found to be more accurate for judgments of higher confidence (4-6) than lower confidence (3-1). This mirrors the first experiment and shows that lower-level FOKs are less accurate at predicting future memory performance than are the higher-level FOKs (1965).

The analyses of the lower confidence ratings in the FOK literature gives rise to many of the theoretical issues that have been presented for JOLs, specifically the dual-process paradigm (Daniels et al., 2009). Current research on memory helps account for the lack of reliability of negative FOKs and its relationship to familiarity. The cue familiarity hypothesis framework introduced by Reder and Ritter (1992) introduces the idea that familiarity can account for memory inaccuracy, especially when it comes to FOKs. A person's prior knowledge of a subject and terms associated with an item can inhibit their memory performance and will cause the person to mistake their own unconscious familiarity with the item's associations for true knowledge of the item. This negative effect of prior knowledge on memory performance has been replicated in both FOK (Carol & Buss, 1988; Carol & Simmington, 1986) and JOL tasks (Toth, Daniels, & Solinger, 2011).

Judgments of learning

The first research to be conducted on judgments of learning (JOLs) was conducted by Arbuckle and Cuddy (1969). Similar to the research by Hart (1965) on FOKs, the first experiment used a simple yes/no procedure to make memory predictions. For this experiment, however, the target items were read aloud and then judged as to whether or not the individuals believed they would remember them at a later test. This dichotomous procedure proved to be effective as perceived memory success (i.e. "yes") was shown to be a significant predictor of future memory performance (Arbuckle & Cuddy, 1969). This procedure was modified in the second experiment by providing a scale of judgments ranging from "very likely" to "very unlikely" as predictors of future memory. Another aspect added was a post-test judgment indicating the subject's confidence that they had produced the correct item, using a judgment scale similar to that used in the study portion of the experiment. The results of the study showed a significant relationship between memory predictions and subsequent performance, particularly with recall (1969).

JOL theory has since progressed and has been the subject of much theoretical investigation over the years (e.g., Koriat, 1997; McCabe & Soderstrom, 2011). These investigations have mostly examined the question of what variables influence JOLs and JOL accuracy. To explain judgments of learning, theoretical approaches have focused on two different subsets of conceptual frameworks: unidimensional and multidimensional (Jang & Nelson, 2005). For unidimensional theories, JOLs and recall are believed to be governed by the same information. At study, then, participants would examine a particular item, assess the strength of their memory for that item, and base their judgments on the level of memory strength they perceived. The *trace-access* theory, for example, postulates that a strong correlation between recall and JOLs is to be expected since they are affected by the same underlying factor (Hart, 1967). Noise, limited degree of precision in the JOL scale, and other variability issues make the correlation between the two less than perfect (Loftus, Oberg, & Dillon, 2004). The *memory-strength* theory, another influential unidimensional theoretical framework, is also concerned with the reasons as to why JOL-recall accuracy is less than perfect (Busey, Tunnicliff, Loftus, & Loftus, 2000). One explanation for this discrepancy is that the fineness of the judgment scale limits the precision with which participants can make their judgments (Jang & Nelson, 2005). Another explanation theorizes that, although the person is trying to selectively assess memory strength from long-term memory, irrelevant information in working memory during the time of immediate judgments adds noise to that judgment, making immediate JOLs less accurate than delayed JOLs (Dunlosky & Nelson, 1992).

In contrast, multidimensional frameworks of JOLs theorize that multiple cues underlie these judgments and drive their accuracy (Jang & Nelson, 2005). The best example of this is Koriat's (1997) *cue-utilization* theory, which suggests that individuals use different cues to predict recall (such as pair relatedness or imageability) and that, in addition, JOLs can be based on subjective experiences such as processing fluency (Koriat & Bjork, 2005). Predictably, processing fluency has been shown to significantly decrease JOL accuracy (see Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Benjamin, Bjork, & Schwartz, 1998; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Koriat & Ma'ayan, 2005; Koriat, 1997; Matvey, Dunlosky, & Guttentag, 2001; Rhodes & Castel, 2008), although the participants' knowledge of the difference between the study and retrieval environments has also been shown to reduce inaccuracy (Koriat & Bjork, 2005).

Similar to the theoretical frameworks associated with JOLs, the methods that are used to analyze the accuracy of these judgments have greatly evolved. Today, JOL accuracy is determined by examining *resolution*, which measures the extent to which individuals can discriminate what will and will not be remembered (McCabe & Sodersom, 2011). High resolution, then, would be an instance of an individual giving high JOLs at study with accurate recall at test or producing low JOLs at study and low recall at test. Low resolution, in contrast, would occur when high JOLs were given for items that were not later remembered and/or low judgments for items that were remembered. One of the most common statistical methods to compute resolution (and one that will be predominantly used in this experiment) is by the use of modified Goodman-Kruskal gamma correlations (Goodman & Kruskal, 1954; Nelson, 1984). A

Gamma correlation is a measure of association (i.e., a correlation coefficient) that is calculated for the individual and then averaged across the sample (Nelson, 1984). In recent studies, gamma correlation coefficients have also been tested using different types of measurement that calculate varying aspects of metamemory. Toth et al. (2011), for example, determined gamma correlations in terms of Standard Coding and Recollection Only Coding. The researchers computed Standard Coding by regarding recollected items that were given R and K judgments as hits while Recollection Only Coding was computed by only regarding recollected items that were judged as R as hits. These two types of coding were used to assess whether JOL accuracy could be enhanced if the analysis was limited to instances of recollection (Toth et al., 2011).

Process dissociation

As previously mentioned, metamemory research, particularly with JOLs, relies heavily on the differences of the quality of the memory for the information that is being produced. For the present metamemory experiment using recall, retrospective judgments (ones that are made after memory production) can be classified into two categories: *remembering*, or the recollection of specific episodic details, and *knowing*, or memory for an item unaccompanied by such details (Gardiner, 1988). Similar to JOLs, these remember-know judgments can also be explained by unidimensional and multidimensional theories, but for the purposes of this research, a dual-state explanation will be used. This dual-state theory holds that these judgments are two perfectly valid, though imperfect, measures of two different states of conscious awareness, and that individuals are experienced in distinguishing between the two (Gardiner, 2001; McCabe, Roediger, McDaniel, & Baltoa, 2009). Many theorists hold that remember/know judgments should be treated as functions of recollection and familiarity (Parks & Yolinas, 2007), though some researchers believe that there has yet to be a definite distinction found between the two types of judgments (McCabe & Soderstom, 2011).

An important paper in our understanding of different memory mechanisms was by Jacoby (1991) in his influential paper on recognition memory. In this paper, he introduced the *process dissociation* framework and suggested that memory is governed by two distinct processes: intentional influences (i.e. recollection) and unconscious influences (i.e. familiarity). In this, he argues that automaticity should not be mistaken for memory of certain details; rather, he states that automatic influences are governed by perceptual characteristics of an item, such as an individual's occupation (Mandler, 1980), and prior knowledge. Jacoby (1991) also suggests that the framework can extend to memory recall and not just recognition. In a more recent review of the process dissociation framework, however, Yonelinas (2002) acknowledges the limitations of the theory to recall due to the different test cues' (i.e. item-cues vs. cued-recall) response requirements and scaling methods. In essence, a scaled drop in accuracy for a recognition task is not proportionate to a drop in accuracy in a recall task (2002). Despite this, Yonelinas has also provided evidence that remember/know judgments are indices of recollection and familiarity (Parks & Yonelinas, 2007).

Immediate vs. delayed

An important development in our understanding of metamemory was the finding that delaying predictive judgments for to-be-remembered items produced significantly higher JOL accuracy compared with those items that were judged immediately. This phenomenon is referred to as the *delayed-JOL effect* by Dunlosky and Nelson (1992). These researchers argued that delayed judgments are based on information retrieved in long-term memory (LTM) and that, in contrast, information that is immediately available in short-term memory (STM) would produce significant amounts of noise in the individual's memory analysis. This *monitoring dual memories* perspective holds that information from LTM is much more indicative of future memory performance than information received from immediately encoding an item (Dunlosky & Nelson, 1992). Another theoretical account of the delayed-JOL effect is the *self-fulfilling prophecy* framework that suggests that judgments made on a delay will be higher for those items that were successfully retrieved at the delayed judgment rather than those that were not (Spellman & Bjork, 1992). In other words, judgments are more accurate for these items since the delay gives the individual an opportunity to make one successful or unsuccessful retrieval before it is required later in the test thus altering the state of the knowledge to be remembered (1992).

Another theory that is believed to govern the accuracy behind the delayed-JOL effect is the *transfer-appropriate monitoring* model, introduced by Dunlosky and Nelson (1997). In the current study, the efficacy of the transfer-appropriate monitoring theory is certainly reasonable since the theory states that the delayed condition provides a similar cued-memory environment to the actual test condition and, in turn, increases individual monitoring accuracy (1997). Some studies have shown, however, that in some instances this framework is not applicable, such as with associative recogntion (see Weaver & Keleman, 2003).

In the present study, we will examine each of these three theories and assess the validity of each as they relate to our results. Theories surrounding the delayed-JOL effect in cued recall tasks for face-name pairs have yet to be examined empirically. Since presentation of the judgment slides shown to the participants after studying the name and face together are identical to the ones shown at test, we expect to find that transfer-appropriate monitoring will provide an accurate explanation to the results. Similarly, the self-fulfilling prophecy framework will also prove to be interesting since the participants will be given a chance to make an initial recall during a delayed judgment slide.

MEMORY FOR NAMES AND FACES

Memory for names and faces is one of the most widely researched areas in cognitive psychology. Although current research has been most prominent with studies on the cognitive differences between younger and older adults (e.g., James, Fogler, & Tauber, 2008; Naveh-Benjamin, Guez, Klib, & Reedy, 2004; Old & Naveh Benjamin, 2011; Tse & Roediger, 2010) as well as in diagnostics in Alzheimer's patients (Greene & Hodges, 1996; Rentz et al., 2011; Werheid & Clare, 2007), it is certainly not limited to this type of research and has been expanded into many different topics in the field. Some of the most intriguing literature exploring these types of associations include event-related potential (ERP) research that explores the different neurological processes that underlie memory for names, memory for faces, and the associations that are made between the two (MacKenzie & Donaldson, 2009; Pickering & Schweinberger, 2003; Weise, Schweinberger, 2008). Early psychological research in this field, however, began much differently and focused more on the two types of stimuli separately, often within the context of the processing (LoP) paradigm originally set out by Craik & Lockhart (1972).

Levels of processing

The LoP paradigm was first introduced to counter the multi-store memory model introduced by Atkinson and Shiffrin (1968). As opposed to thinking of memory as a storage system that can be broken down into sub-units (sensory memory, short-term memory, and long-term memory), the LoP paradigm introduced a more dynamic model that suggested that memory performance is a function of depth of mental processing (Craik & Lockhart, 1972). This depth of processing is seen as a continuum that ranges from shallow to deep processing with intermediate values in between. Shallow processing (e.g. processing based on the letters of a name) has been shown to yield a weak trace of that memory in storage, one that is susceptible to forgetfulness when the trace decays. Deep processing (e.g. processing based on meaning), however, has been shown to leave very strong traces in memory and greatly decreases the probability of trace decay even after longer periods of time. As the paradigm has matured, other variables such as transfer-appropriate processing and familiarity have been found to be appropriate modifiers of the theory (Lockhart, 2002; Toth, 1996).

In researching memory for faces and names, the levels-of-processing theory became significant as its application to semantic encoding was explored. In one study exploring memory for personal information, McWeeny et al. (1987) ran an experiment that prompted the participants with the last name and occupation of 16 given faces and asked them to recall the information when prompted with the corresponding face. Any contextual graphics that could have provided clues as to the faces occupations were removed and each face was randomly assigned an occupation and surname. The results from the study showed that names of faces were significantly harder to remember than were the occupations assigned to them, even when the occupation was “baker” and the last name “Baker” (1987). This implicates semantic encoding (i.e. deep processing) as a contributive memory mechanism in the study and further suggests that information that has conceptual meaning will be processed deeper than information of a more arbitrary nature. In remembering name-face pairs, it is now common practice to encourage building semantic relationships between the face stimuli and the corresponding name (Wiese, 2011), particularly in memory research with older adults (Schmidt, Dijkstra, Berg, & Deelman, 1999; Troyer, Haefliger, Cadieux, & Craik, 2006; Yesavage, Rose, & Bower, 1983).

Model of facial recognition

Advances in levels-of-processing research performed on name-face associations and semantic memory lent itself to the model of facial recognition originally set out by Bruce and Young (1986).Their model postulated that different sub-processes worked in unison to encode faces into memory. When an individual first views a face, a "view-centered description" is derived from the perceptual input and records the features that are important in facial recognition, such as age and sex. As this initial input is processed, the information that it derives provides a basic facial structure that is used to help compare that face to others that are stored in memory. After the initial phase, the structure-based information is then transferred to "face recognition units" that are used in conjunction with "personal identity nodes" to help identify a person from their memory (1986). Researchers believe that this facial-recognition process is governed by the fusiform gyrus in the temporal lobe (Kanwisher, McDermott, & Chun, 1997; Kanwisher & Yovel, 2006).

In this model of facial recognition, semantic memory has been shown to increase the strength of the "personal identity nodes" for the face and, subsequently, the memory recall of an individual's personal information (Bruce, 2009). Deep encoding of both facial and personal characteristics, then, will lead to better production of a name when prompted with a face.

Intention to learn

Another important aspect of memory for names and faces that has been shown to significantly affect performance is the intention to learn. First introduced in the academic literature by Kausler and Lair (1965), intentional learning refers to the idea that participants will yield much greater recall for items when they are told that their memory will be tested at a later time. This original study on intentional learning and paired-associates has recently been modified to test memory for names and faces, most notably by Troyer et al. (2006). In their experiment, intentional learning, along with self-generated information for the names and faces (i.e. semantic memory), yielded the highest memory accuracy for free recall of names, yes-no recognition of faces, cued recall of names, and name-face matching in older adults (Troyer et al., 2006).

Metacognition and name-face associations

Mandler (1980), in his famous paper outlining theories associated with familiarity in recognition memory, first helped describe metacognition and unconscious memory for names and faces by providing an example from a bus ride. In this example, an individual spots a man whom he is certain he has seen before, but he is not sure how or from where or when he knew him. As the individual searches through his memory, he sifts through semantic information (i.e. if he was seen at the supermarket) to help procure his association with the man. Eventually, the man realizes that the man is somebody that he sees regularly ("That's the butcher from the supermarket!") (p. 253). Today, this example is commonly referred to as the "butcher on the bus" phenomenon in which a relatively automatic form of memory (familiarity) encourages a search for recollective details about a target item. The phenomenon has also laid a basic foundation for current research in memory and metamemory for name-face associations.

Following Mandler's study, research on metamemory for names and faces still remains a fairly new line of inquiry. Separately, though, these two types of stimuli have garnered many interesting projects researching metacognition. In researching Jacoby's (1991) process dissociation framework, for example, Parkin, Gardiner, and Rosser (1995) examined how different types of rehearsal would affect recollection and familiarity when studying faces. They found that when the faces were studied under conditions of divided attention, the number of “remember” responses significantly decreased while the amount of “know” responses remained relatively unchanged. In the second part of their study, participants were prompted to study faces that were shown either immediately or after a delay. In this, the “spaced” condition increased “remember” responses and reduced “know” responses. This research on facial recollection and metacognition has also been extended to facial distinctiveness and differing contexts (Godden & Baddeley, 1980; Sommer, Heinz, Leuthold, Matt, & Schweinberger, 1995).

Metacognitive research exploring memory accuracy for names and faces together still remains relatively untouched. Much like the current experiment, recent studies on the subject have been interested in the unreliable nature of name and face learning. Tauber and Rhodes (2010), in researching the metacognitive errors that lead to memory inaccuracy for proper names, found that JOL accuracy was much higher for common information attributed to a person (i.e. occupation) than was the proper information (i.e. name). These results reveal a need for provisions when it comes to remembering name-face combinations due to the metacognitive deficits that come with them (2010).

In researching metacognitive deficits in older adults, Toth, Daniels, and Solinger (2011) tested memory accuracy for both young and older adults using the names and faces of actors from the 1950’s and 1990’s. In this, they found that prior knowledge of the 1950’s actors helped older adults in remembering episodic details about studying the actor's name, but also decreased their JOL accuracy for the same actors by increasing feelings of familiarity which, in turn, distorted their predictive memory judgments. This research also implicates recollection as the driving force behind metacogntive accuracy (2011).

Some interesting cognitive research has also been dedicated to the neural correlates of metamemory and JOLs using memory for names and faces. Do Lam et al. (2012) used name-face pairs to determine if the neurocognitive mechanisms that govern metamemory strategies (the medial prefrontal cortex) are linked with the mechanisms that govern memory and retrieval (the medial temporal lobe). Overall, they found that JOLs incorporate retrieval processes in memory and metacognitive accuracy for names and faces.

OVERVIEW OF THE PRESENT STUDY

The goal of the present study was to examine memory and metamemory for face-name associations using a dual process approach to JOLs (Daniels et al, 2009; Toth et al., 2011). Participants studied a set of 96 face-name pairs, making JOLs to each pair based on the face and the first letter of the associated name. Half of the JOLs were made immediately after first studying a face-name pair, while the other half were made after a delay of six or seven intervening face-name pairs. Then, at test, participants were shown the faces along with the first letter of the corresponding names and asked to recall the names that were paired with the faces. When a name was produced, participants were further asked to describe whether they recollected the name as being paired with the face ("Recollect"), whether the name came to mind automatically without any recollective details ("Know"), or whether they had no memory for the face-name pairing ("No Memory").

We predicted that name recall would be better in the delayed, as compared to immediate, JOL condition based on the idea that a delayed JOL trial provides participants with a second study opportunity. This result would also be consistent with the *self-fulfilling prophesy* hypothesis of the delayed JOL effect (Spellman & Bjork, 1992). We also predicted a delayed JOL effect with our face-name stimuli, such that JOL accuracy would be higher in the delayed, as compared to immediate, JOL condition (Watier & Collin, 2011). Finally, replicating prior work by Daniels and colleagues (Daniels et al., 2009, Toth et al., 2011) we expected JOLs to be highest for names that were actively recollected (i.e., given a "Recollect" judgment) at test, compared to those that were generated more automatically, accompanied only by familiarity ("Know") or no feeling of memory ("No memory"). Importantly, however, we expected the pattern of JOLs for Recollect, Know, and No Memory judgments to differ in the immediate and delayed conditions, reflecting the differential diagnosticity of these judgment at these separate points in time. This pattern would be consistent with the *monitoring dual-memory hypothesis* (Dunlosky & Nelson, 1992). It would also extend Daniels' et al. dual-process approach to JOL accuracy by showing that, rather than being exclusively based on recollection, under the appropriate conditions JOL accuracy could also be supported by more automatic forms of memory.

METHODS

Participants

Fifty-two young adults (n_{female} = 36, n_{male} = 16; mean age = 19.21) were recruited for the experiment. All of the participants were undergraduates at the University of North Carolina Wilmington and participated in the experiment in exchange for course credit. The data from four subjects were not included in the final results; two due to computer failure, one for being underage (17 years old), and one for failing to answer all but nine of the 144 test items. The final sample consisted of 48 young adults (n_{female} = 34, n_{male} = 14; mean age = 19.15).

Materials

The stimuli consisted of names and faces obtained from the Face Database created by Park and Minear (2004). One-hundred and forty-four neutral Caucasian faces were chosen from the database with equal numbers of men and women. Out of the database's list of faces, we discarded pictures with features that we considered too distinctive, including unusual hair styles or color, facial piercings, distracting jewelry, and overly colorful clothing. The faces that were not discarded were then divided into two age groups: young (18-29 years of age) and old (60+ years of age). The stimuli were then categorized by gender, leaving four groups of 144 faces: 36 young-female, 36 young-male, 36 old-female, and 36 old-male. Next, the names that would correspond to the faces were selected from the US Social Security Administration's online name database (2011). The most popular names for newborns from the years 1940 and 1990 were used to represent the old and young faces, respectively. These names were controlled for overlapping stimuli (names occurring in both lists) then randomly assigned to the facial stimuli in their respective demographics.

After assigning the names to the faces, the 36 stimuli in each of the four demographics were divided again and assigned to three groups of 48 names and faces (12 of each demographic). These three groups of pairs were rotated in three counter-balances so that each participant would study and be tested on one set of name-face pairs with immediate JOLs, another set of name-face pairs with delayed JOLs, and another set of new/unstudied stimuli that were only presented in the test phase. The experiments were programmed and run using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Eleven keys were labeled across the top of the keyboard, "~" to "0", but were re-labeled "0" to "100" in intervals of 10 in order for the participant to make predictive memory estimates.

Procedure

Each participant was tested in a separate room with a computer and a research assistant to facilitate. Prior to beginning the experiment, each participant was asked to fill out a brief survey as well as a consent form approved by UNCW's Institutional Review Board (IRB). If the participant was under the age of 18, they were asked to fill out a separate consent form and were allowed to take part in the experiment to gain credit, though their results were labeled as unusable. The experiment began with instructions on how the participant would be shown two kinds of slides during the study phase, one showing a face-name pair and a second that would ask the participant for a judgment of how well they would recall the name when later shown the face and first letter of the name. The directions also explained how half of the name-face pairs would have judgment slides that would be delayed by six or seven other study trials that were both immediate and delayed. After completing four practice items, the subjects were then asked to give a global JOL for how well they would recall the 96 names at test.

During the study phase, participants saw a randomized list of 48 immediate items and 48 delayed items. Although the order in which the pairs were presented was randomized, the manner in which the delayed and immediate items were presented was manually programmed by the researchers prior to the experiment in order for the two slides in each delayed item to be correctly separated by a certain number of stimuli. To make sure that all of the delayed items would have equal delays, filler items were added at the end of the study phase. These stimuli were not used at test.

In the judgment trials, the participant was presented with a JOL scale (0 to 100) that would indicate how likely they were to recall the individual's name when presented with the corresponding face. A rating of 0 indicated that the participant would absolutely not be able to remember the name on a future test while a rating of 100 would indicate they would be absolutely certain they would be able to remember the name. The participant was encouraged to use any reasonable means to remember the faces, including facial features and mnemonic devices. Participants were also asked to use the entire JOL scale to present their levels of certainty more accurately. The participant did not have a time limit to study the pair and make their JOL decision. After entering the JOL, the next face-name pair would immediately appear on the screen.

Upon completion of the study phase, participants were asked to give another global judgment indicating how well they believed they would be able to remember the names of the 96 stimuli. This judgment represented the relative percentage of names they would be able to correctly recall when shown a face and the first letter of the name. After this judgment, the participants were given instructions on how to complete the test portion of the experiment. These directions explained the nature of the recall task and which memory judgments corresponded with certain types of memory. The participants were asked to try and recall the name that was paired with the face at study. If an individual was prompted with a face and the first letter of the name and remembered the corresponding name along with specific details about the pair (such as facial features, hair, etc.), he/she would tell the experimenter the name that they recalled and a "Recollect" judgment. If the subject remembered the name when prompted with the corresponding stimuli but did not specifically remember details associated with initially studying it, then the participant would be asked to provide the name and a "Know" judgment. Like the definitions originally set out by Toth, Daniels, and Solinger (2011), this term was substituted for "Familiarity" so that the subjects would not confuse familiarity for the face in the pair with the strength of their memory for the name (see McCabe & Soderstrom, 2011, for more on the subject). If the participant had no memory for the item or could not remember the name that corresponded with the face, then the subject was asked to respond with "No Memory" and the name that they believed corresponded to the item. If the individual could not remember the name and could not produce a name that could belong to the face, then he/she was asked to respond with “Pass” and the response was coded with five “x’s”. Each of these stimuli was presented one at a time with no time limitations. Similar to the study phase, each item had two slides—one that presented the face with the first letter of the name and another that showed the face, the name that was recalled by the participant, and the judgment options “Recollect,” “Know,” and “No Memory” at the bottom of the screen. For this task, the participant verbally told the experimenter the name that they recalled and the memory judgment associated with the recalled name for the experimenter to type in to the computer. This was done in order to reduce the number of different possible spellings for a given name so that the memory data could be coded more easily.

After completing the recall task, the participants were thanked, given course credit, and dismissed from the experiment.

RESULTS

For the analyses of the results, t-tests and ANOVAs with a repeated-measures design were used to explore possible significant values. Each significant t-test reported includes the t-critical value (*t*), the degrees of freedom (*df*), and the *p*-value. For ANOVA’s with significant values, the *F* value, the *p*-value, mean square error (*MSE*), and effect size (partial eta squared, or η_p^2) were reported. Unless otherwise specified, an alpha level of .05 ($p < .05$) was used for all statistical tests.

Judgments of learning

To begin the analysis, we examined the average judgments that participants made throughout the experiment as functions of the types of ratings, the conditions the judgments were made in, and of their overall performance. Prior to and immediately after studying the name-face items, participants were prompted to make a global judgment indicating the proportion of items

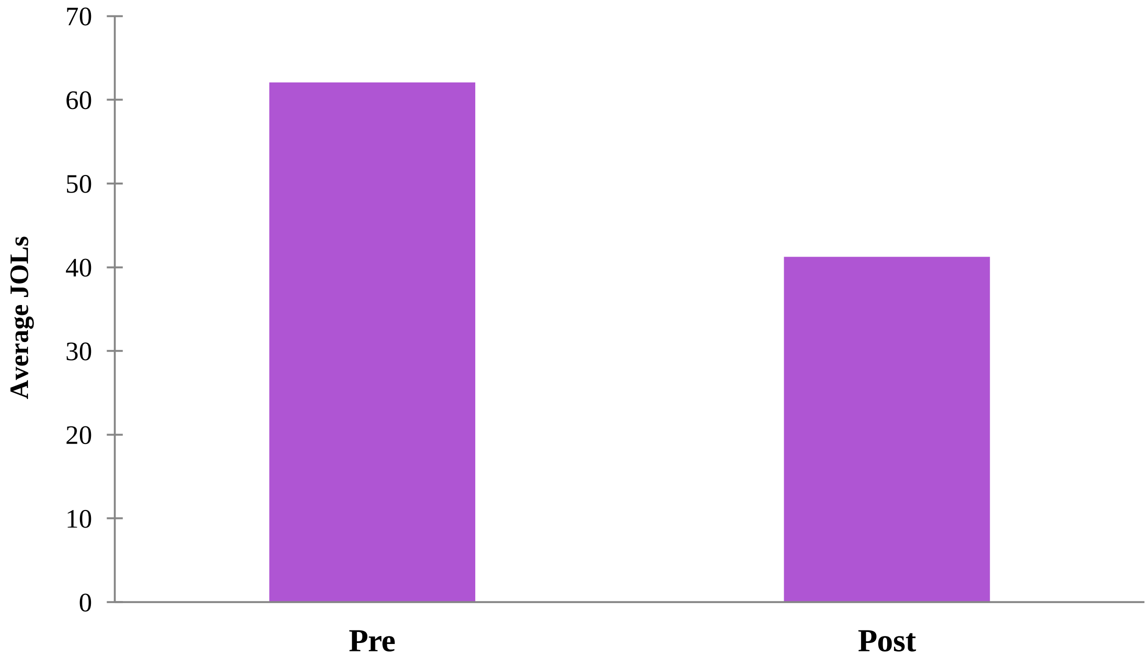

Figure 2. Mean JOLs provided by the participants before and after the presentation of the name-face items during the study phase of the experiment. A t-test showed a significant decrease in predicted memory performance after studying all of the pairs.

they believed they would remember at test. A paired-samples t-test analyzing these JOLs indicates a significant decrease in average judgments between the pre- and post-study JOLs, $t(47) = 8.90$, $p < .001$. This shows that, overall, participants tended to lower their expectations about future memory performance after all of the name-face pairs were presented (Figure 2).

A similar analysis was performed on the list-wide (LW) JOLs for both the delayed and immediate conditions. These LW JOLs are calculations of the averages of all the JOLs made in every judgment type (R/K/N) and provide a holistic perspective to the JOLs made in each condition. When the LW JOLs were analyzed across conditions (delayed vs. immediate) using a paired-samples t-test, no effect was found, $t(47) = .817$, $p < .417$. While this pattern initially seems counter-intuitive to the theoretical expectations set out by the delayed-JOL effect, the lack of effect can be explained by examining the interaction amongst the average JOLs made for R, K, and N judgments by condition (see *JOL Accuracy*, below).

Lastly, we examined the interaction between the immediate/delayed conditions and the test responses and found the relationship to be significant, $F(2, 94) = 46.107$, $MSE = 91.727$, $p < .001$, $\eta_p^2 = .495$. This shows that *Recollect*, *Know*, and *No Memory* each have significantly different values from one another when analyzed as a function of their JOL type (see Figure 3). We wanted to explore this interaction further and ran t-tests on the average JOL scores for each response type. We found that the average JOL for R responses in the delayed condition was higher than that for the immediate condition, $t(47) = .\text{-}5.908$, $p < .001$, while that for N responses was lower, $t(47) = 6.358$, $p < .001$. There was no significant difference between JOLs for K responses in the immediate versus delayed conditions, $t(47) = -.19$, $p = .85$. The data thus

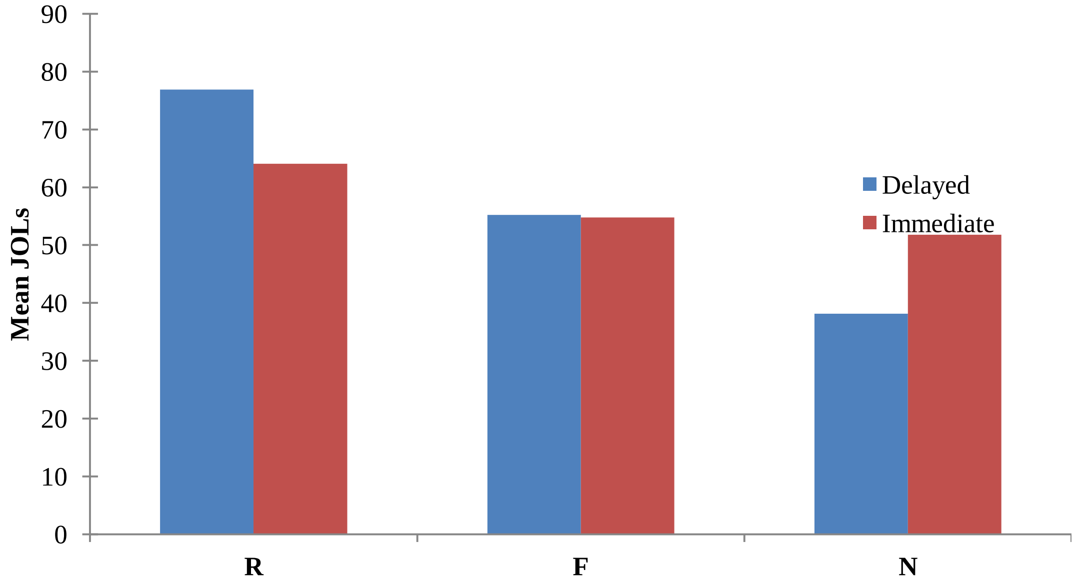

Figure 3. Mean JOLs for different response types (R/K/N) as a function of their condition (delayed/immediate).

show an interesting cross-over pattern in which the delayed R judgments are higher than the immediate R judgments, the K judgments almost identical in each condition, and the delayed N judgments are lower than the immediate N judgments. The average JOLs are higher in both conditions for conscious influences on memory, but for Know and No Memory responses in the immediate condition, the judgments are not significantly different. Compared to judgments made in the immediate condition, whose average JOLs decrease across the responses, those made immediately are much more prone to inaccuracies in perceived future performance for items that are later recalled with lower levels of certainty (i.e. Know and No Memory).

Memory performance

To calculate memory performance, we examined the average JOLs made for each R/K/N judgment and then took the difference of the "hit" rates (i.e., R & K judgments to studied names that were recalled) and "false alarm" rates (i.e., R & K judgments to unstudied names that were output). Table 1 shows the proportion of R, K, and N responses for both studied and unstudied stimuli. Note that there was no immediate/delayed distinction for unstudied stimuli (i.e., a common set of face+first-letter cues was used to establish a false alarm rate for both conditions) so this data is reported in the "Overall" row only. Note as well that the table includes a measure of estimated familiarity (F) equal to K/(1-R) based on the assumption that Recollect and Know are based on familiarity and recollection which make independent contributions to memory performance (see Jacoby, Yonelinas, & Jennings, 1997). We also looked at the overall mean accuracy of the scores by subtracting the proportion of false alarms from the proportion of hits of recollect (R), know (K), and computed familiarity (K/[1-R]) (Figure 4).

		Studied				Unstudied			
		R	K	K/(1-R)	N	R	K	K/(1-R)	N
Delayed	*M*	**0.31**	**0.34**	**0.51**	**0.35**				
	SD	0.17	0.18	0.24	0.21				
Immediate	*M*	**0.25**	**0.31**	**0.42**	**0.45**				
	SD	0.17	0.19	0.24	0.22				
Overall	*M*	**0.28**	**0.32**	**0.46**	**0.40**	**0.03**	**0.15**	**0.15**	**0.83**
	SD	0.16	0.18	0.23	0.21	0.05	0.18	0.19	0.19

Table 1. Proportions of R, K, K/(1-R), and N as a function of condition (immediate/delayed) and whether they were studied or not. All of the Unstudied proportions are similar since those items were only presented at test.

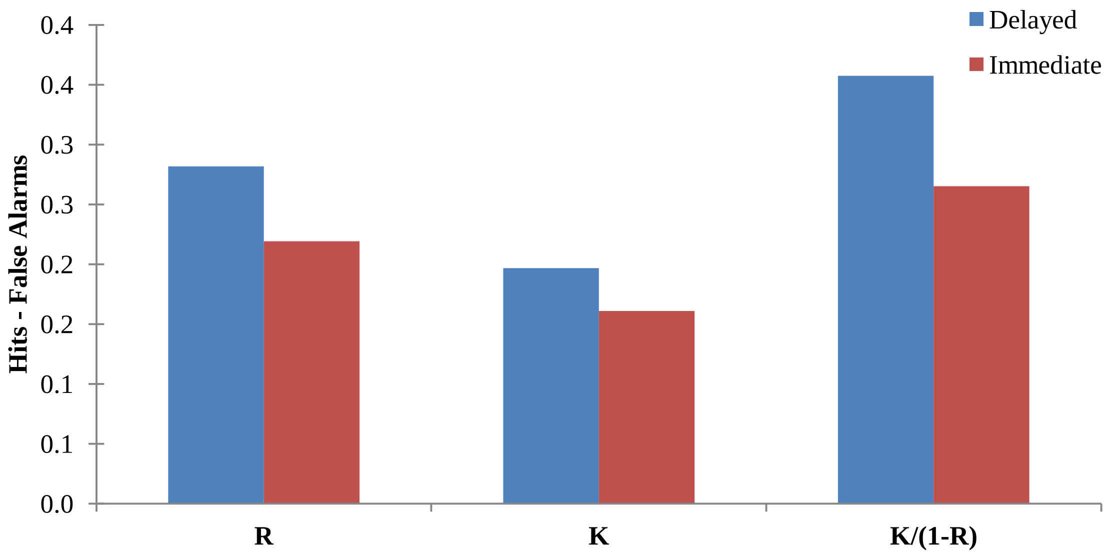

Figure 4. Overall memory accuracy for the studied items as a function of response type (R/K/N) and condition (immediate/delayed) with computed familiarity (K/[1-R]).

The analysis of the accuracy of overall memory performance (i.e. hits – false alarms) showed a significant difference between the immediate and delayed conditions, $F(1, 47) = 39.277$, $MSE = .007$, $p <.001$, $\eta_p^2 = .455$. No matter which judgment was made at test, the condition seems to enhance overall memory performance, a result that we expected to find as per the expectations of the delayed-JOL effect. The analysis also showed a significant difference in accuracy scores of overall memory performance for the different response types (including calculated familiarity) regardless of the condition, $F(2, 94) = 18.271$, $MSE = .023$, $p <.001$, $\eta_p^2 = .280$. Recollect, know, and calculated familiarity were all found to be significantly different from each other with R having more overall accuracy than K, but lower than F. This higher overall memory accuracy for computed familiarity was also found in the results of Toth et al. (2011).

Lastly, we found significant interactions between the JOL type and condition for overall accuracy in memory performance. To explore this, three paired-samples t-tests were run comparing the immediate and delayed conditions in each judgment type. With the R judgment, a significant difference in the delayed and immediate conditions was found, indicating higher corrected memory performance for items that were delayed, $t(47) = -3.937$, $p < .001$. This difference between the immediate and delayed conditions was also found for the K judgment, $t(47) = -2.536$, $p < .01$, as well as calculated familiarity (F), $t(47) = -5.241$, $p < .001$. This indicates that conscious (R), automatic (K), and calculated familiarity (F) are different from one another as well as when they are applied as factors of the immediate and delayed conditions, though the significance is not as strong as the other *F*-values. Overall, this indicates that while the memory accuracies amongst the judgment types and between the conditions are significantly different, these two variables together do not produce largely different memory accuracy values, though the relationship is significant

JOL accuracy

The very last analysis that we made with these data was to examine the overall JOL accuracy according to condition (immediate/delayed). To do this, we used Goodman-Kruskal (1954) gamma correlation coefficients modified for metamemory studies. In our study, gamma correlations were computed by dividing the total number of concordant memory judgment pairs (P) minus the discordant pairs (Q) by the sum of the concordant and discordant pairs together, or gammas = (P-Q)/(P+Q). Every subject's gamma correlation coefficient was then averaged together according to condition. To differentiate the different ways memory can be defined, three different calculations were used: a standard recognition computation that regards both R and K responses as "old" judgments and N responses as "New" judgments, a "recollect only" computation that treated R responses as "old" and K and N responses as "new", and a recall only computation that simply determined whether the target item (i.e. the name originally paired with the face) was produced. With each of these different calculations, the delayed and immediate conditions were differentiated, as shown in Figure 5.

Overall, gamma correlations for delayed judgments in each of the three codings were found to be significantly higher than those made in the immediate condition. To explore these interactions, we ran paired-sample t-tests comparing the delayed and immediate gamma correlation coefficients in each calculation. Gammas in the Standard Coding show that metamemory was significantly more accurate when the JOLs were delayed, $t(47) = -2.152$, $p <$ 05. Accuracy was also found to be significantly increased for delayed JOLs in Recall Only Coding, $t(47) = -3.536$, $p < .001$. With Recollection Only Coding, however, the two conditions

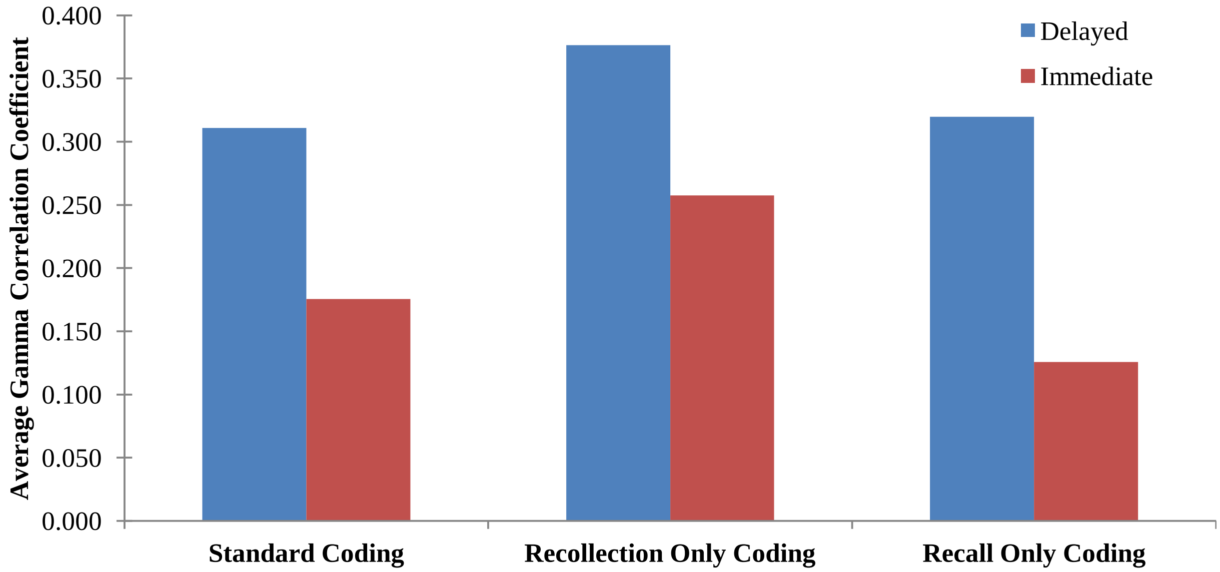

Figure 5. Average Goodman-Kruskal gamma correlations for the delayed and immediate conditions as a function of memory condition coding. Standard Coding was computed by regarding the R and K judgments as "old" and N as "new", Recollection Only Coding was computed by only regarding R as "old" and K and N as "new", and Recall Only Coding was computed by scoring data based on whether the participant produced the correct target.

were not found to be significantly different from each other, $t(47) = -1.96$, $p = .056$. Furthermore, the analysis of the variability of the accuracy of the immediate condition in all codings was found to be much higher than in the delayed condition, indicating both a jump in JOL accuracy when the judgments are delayed and lower reliability in JOL accuracy for the immediate condition. We also found that the average gamma correlations in each calculation (Standard, Recognition Only, Recall Only) were significantly different from each other, $F(2, 94) = 3.848$, $MSE = .061$, $p < .005$, $\eta_p^2 = .076$. Lastly, we found that the average gamma correlation coefficients for each coding calculation as a function of their condition were not significantly different from each other, indicating a consistency in gamma correlation coefficients for each condition over the codings, $F(2, 94) = .748$, $MSE = .050$, $p = .476$, $\eta_p^2 = .016$. This homogeneity in the gamma correlations confirms the accuracy of the calculations since they are expected to compute relatively similar values despite their conceptual differences.

To examine the relationship between mean JOL (delayed vs. immediate) and response (R/K/N) type, a repeated measures, within-subjects ANOVA was computed on the JOLs, sorted according to how the relevant pairs were classified in the test phase (R/K/N). Given the results from the previous List-Wide analysis, we did not expected to find a main effect for the immediate versus delayed conditions, which indeed was what we found, $F(1, 47) = .005$, $MSE = 191.801$, $p = .943$, $\eta_p^2 < .001$. Again, this indicates that overall, participants did not report higher JOL in the delayed versus immediate condition. However, we did find a main effect for the three types of responses (R/K/N) at test, $F(2, 94) = 98.46$, $MSE = 161.01$, $p < .001$, $\eta_p^2 = .677$. Regardless of the condition, the means of the JOLs for each response type were found to be statistically different from one another. As expected, average JOLs for Recollect was significantly higher than those for No Memory, $t(47) = 11.298$, $p < .001$. Average JOLs for

“Recollect” judgments were also significantly higher than those for “Know” judgments, $t(47) = 9.040$, $p < .001$. Finally, “Know” judgments had significantly higher JOLs in comparison to “No Memory” judgments, $t(47) = 5.632$, $p < .001$. This indicates that when participants reported a “Recollect” judgment at test, the JOLs they provided for the item were significantly higher than those given to items rated as “Know” or “No Memory” at test, results that we expected to coincide with the process-dissociation framework for metamemory (Gardiner, 1988).

GENERAL DISCUSSION

In reviewing the results of this study, we will focus on the main purpose of this research, which is to explore the underlying processes that govern our accuracy (or inaccuracy) in memory for names and faces. First, the influences of recollection and familiarity on metamemory predictions will be examined using the framework introduced by Daniels et al. (2008) and later appended by Toth et al. (2011). Their research concluded that recollection memory and memory accuracy is governed by controlled (recollection) and automatic (familiarity) influences and that both bring significant contributions to the table; however, in some cases, the latter can negatively affect predictive memory accuracy. In testing JOL-accuracy in memory for famous celebrities from the 1950’s and 1990’s, they found that older adults would give erroneously higher JOLs for actors that they were familiar with (i.e. from the 1950’s) than for the actors that they were not familiar with (i.e. from the 1990’s). Prior knowledge, an established influence on familiarity (see Reder & Ritter, 1992), was found to be the root cause of the older subjects’ inaccuracies (Toth et al., 2011). In regards to familiarity, we modified the dual-process memory scheme for recollection used by Daniels et al. (2008) to encompass recall memory by replacing the Recollect/Familiar judgments from the prior research with Recollect/Know judgments, assuming that they are the most accurate recall-based correlates that have been demonstrated in the academic literature (McCabe & Soderstrom, 2011). Overall, the results do suggest that familiarity is much less accurate in determining future memory performance, particularly with back-sorted JOLs in the immediate condition.

Furthermore, this research also sought to explore the influence of judgment timing on JOL-accuracy. In the metamemory literature, the delayed-JOL effect has been proven to be a strong, positive influence on predictive and prospective memory accuracy (see Nelson & Dunlosky, 1991; Dunlosky & Nelson, 1992). While this effect has also been observed with cued recall involving name-face associations (Watier & Collin, 2011), we also wanted to explore the dual-process mechanisms that are affected by delaying judgments, particularly with automatic influences on memory (i.e. know). Overall, allowing participants to make JOLs after a delay improved predictions in future memory performance in each analysis of the data. Furthermore, controlled memory influences (i.e. recollection) benefitted greatly from the prolonged timing of the judgments while JOL accuracy in regards to automatic influences did not show any benefit and, in some cases, decreased. These results, along with the ones discussed in the earlier paragraph, will be discussed in the context of the dual-process approach to JOL accuracy proposed by Daniels et al. (2009) as well as the prior literature discussing the delayed-JOL effect.

Recollect/know judgments as influences on metamemory accuracy

Consistent with prior research, controlled influences on memory (R) were found to be much more accurate predictors of JOL accuracy than were automatic influences on memory (K). These results are most evident in the average JOLs back-sorted by judgment type and corrected memory performance calculations. From the data presented, we can conclude that name-face pairs that are encoded in either of the memory conditions (immediate or delayed) will show greater overall accuracy for items that were produced at test and subsequently given R ratings. Conversely, we can also conclude that names produced at test and subsequently given K ratings will have significantly lower accuracy than ones given R ratings, though there are no significant differences in the delayed and immediate conditions for back-sorted JOLs, a result that will be discussed in the subsequent section. These results reflect the process-dissociation framework set out in prior research (Daniels et al., 2008; Toth et al., 2011) and expand our understanding of conscious and unconscious memory to name-face associations.

In exploring the theoretical implications of this dissociation, many different hypotheses are available to explain memory for names and faces. Wixted and Mickes (2010), for example, provided a single-state account for accuracy in prospective memory judgments by viewing memory as a continuous variable determined by the strength of the memory and not by content. Indeed, it would seem likely that R/K/N judgments for name-face associations would be driven by the strength of the memory for the item (i.e. *how well did I remember his/her name?*), but in terms of operational and reliable delineations of memory strength, this framework can be troublesome. Our definitions of R and K revolved around this issue since we believe that continuous memory strength for name-face pairs can be easily misinterpreted. A good example of this is mistaking distinctiveness of a face (the memory cue) for increased memory strength for the name (the target) and the pair as a whole. In testing metamemory for name-face pairs, participant knowledge of what is truly important (or memory for the name when presented with the face) is key in collecting and analyzing accurate predictions of future memory.

Instead of theories that emphasize strength-driven judgments, we argue that the results of this experiment favor multi-dimensional models that are driven by content rather than perceived memory strength. According to these multi-dimensional theories, R and K are two separate processes that reflect different memory experiences and are distinguished based on the amount and quality of information that can be retrieved at the production (or lack thereof) of an item (McCabe & Soderstrom, 2011; Rajaram, 1998). In this framework, participants are also skilled monitors of their own memory and effectively distinguish between the two processes in their own lives (McCabe & Geraci, 2009). We encouraged participants to use this innate expertise by defining R as a state of memory in which specific details are remembered that link the name and face together, such as facial oddities, hair types, and other distinctive features. This reflects common experience since individuals rely on facial distinctiveness to provide a memory cue that allows for them to produce the name that was associated with the face. Conversely, we defined K as being a state of memory in which an individual believes that he or she "knows" the name that is associated with the face, but cannot produce any distinctive details that would tie the two together. This is also emulates casual memory experiences and reflects the common phenomenon of "knowing" a person's name, but not specifically remembering how or why their name was associated with their face.

Now that a multi-dimensional model of memory judgments (R/K/N) has been established, the issue now becomes how controlled and unconscious aspects of memory differ. Generally, automatic influences on memory are believed to be a more form of memory since they are shaped by heuristics such as availability of information and processing fluency (see Jacoby, 1991). We believe that this is the case for the current experiment and draw upon the experience of examining a face and name together. What is immediately available to the participant at the time of the judgment, such as a name and a face, is not an accurate indication of what will be remembered at a later time as it draws on unconsciously processed information that is either lost or misinterpreted at test (Toth et al., 2011). In memory for these pairs, we can also see that prior experience in processing names and faces can contribute to automaticity and memory inaccuracy since the process is erroneously believed to be fast, automatic, and effortless (2011; Jacoby, 1998). This effect can be observed in the stark difference between global JOLs that were given by the participants before and after the study phase. The 20-point decrease in confidence indicates a tendency for individuals to perceive their expertise in memory for names and faces to be much greater than it actually is.

Another possible explanation for the dissociation in results that we found is distinctiveness in the faces and names. In a recent study, Watier and Collin (2012) found that facial distinctiveness significantly increases the probability of recalling the name while faces that were deemed as more “normal” showed lower levels of memory accuracy. While our experiment was not designed to test the effects of distinctiveness, identifying the faces that were found to be the most distinctive and back-sorting the JOLs and memory judgments for them would be an interesting addition to our research.

Judgment timing as an influence on metamemory accuracy

The second effect that was expected to be found in our results was the delayed-JOL effect in metamemory and JOLs (Nelson & Dunlosky, 1991; Dunlosky & Nelson, 1992). Consistent with our predictions, metamemory accuracy for names and faces was significantly increased when judgments were delayed by six or seven other stimuli. This is most evident in all of the gamma correlation coefficient calculations, the back-sorted JOLs, and the corrected memory performance data. Overall, these results indicate decreased accuracy in the immediate condition across the judgment types, with the exception of Know in the back-sorted JOLs.

Similar to the interpretation of the dual-process framework that was discussed earlier, the delayed-JOL effect has several different theoretical frameworks to explain the phenomenon.

Dunlosky and Nelson (1997) have hypothesized that the delayed-JOL effect is a result of transfer-appropriate monitoring (TAM), or that the environment in which the delayed judgment is made mimics the cued-recall environment at test, therefore allowing the subject the opportunity to review the exact environment that they will experience later at test. In our study, the delayed and recall environments were almost identical to one another as they showed the face and the first letter of the corresponding name, so it is plausible that the similarities can account for the increased accuracy. Previous research has indicated, though, that TAM does not account for memory performance with respect to names and faces, specifically with repetition priming (Burton, Kelly, & Bruce, 1998).

Beyond TAM, there are two remaining frameworks that could account for the delayed-JOL effect in name-face associations. In their original research, Dunlosky and Nelson (1992) theorized that lower accuracy with immediate JOLs is a result of faulty monitoring since the individuals making the judgments only have what is readily available in STM on which to base their decision. With delayed JOLs, however, the accuracy is increased since the individuals making the judgments can base their decisions off of what is retrieved in LTM, a much more accurate diagnostic feature of future memory performance since LTM will ultimately determine recall (1992). This monitoring dual memories perspective does seem plausible for our experiment since the participants judged how well they had encoded the names when cued with the corresponding faces when forced to make delayed JOLs. An alternative explanation to this, however, is the self-fulfilling prophecy theory introduced by Spellman and Bjork (1992). When the participants made a judgment for a name when cued with the corresponding face, they had the opportunity to assess their memory for that name before encountering the same cued-memory environment later at test. If the name could be produced, then a higher and more accurate rating would be given; if not, then a lower rating would be given, but with the same accuracy since both instances made successful memory predictions.

One of the more interesting results of the study was the interaction between the time of judgment (immediate or delayed JOLs) and judgment type (R/K/N) with the back-sorted JOL data. In this, the average JOLs for the delayed condition made a predictable stair-stepping sequence in which the Recollect judgments show the highest average JOLs, the Know judgments show lower but medium average JOLs, and the No Memory judgments show the lowest average JOLs. This is an expected effect since conscious influences on memory have shown to produce much higher JOLs and memory accuracy (Daniels et al., 2009). In contrast, automatic influences, while providing some positive effects on memory, have been shown to be much less accurate predictors of future memory performance and, in some cases, been shown to significantly hinder metamemory accuracy (Toth et al., 2011). Participants who produce names for their respective faces without remembering any specific details associating the two should provide judgments that are indicative of a level of uncertainty. Finally, No Memory judgments should be the lowest since, a result that was observed in the delayed condition. In the immediate condition, however, the results are somewhat different. While R judgments did garner higher average JOLs in the immediate condition, there is almost no difference in mean JOLs between the K and N judgments. Participants were as likely to give similar predictive ratings to names that were either produced as a product of automaticity or not produced at all. One possible explanation for this effect is the cue-utilization framework that suggests that inaccuracy in immediate JOLs can be explained by the judgments being based exclusively on components of the subjective (automatic) experience at encoding, specifically processing fluency (Koriat, 1997; Koriat, Bjork, Sheffer, & Bar, 2004). According to this framework, the decreased JOLs of R in the immediate condition when compared to the delayed condition and the similarities in K and N in the immediate condition are a result of participants erroneously judging their on-line encoding experience with names and faces as sustainable memory for the pairs. This would cause items that would eventually be forgotten to be given ratings that were as high as those for names that were eventually judged as K and leave no clear distinction between the two ratings.

Conclusions

Overall, the implications of our findings mark new territory for the understanding of memory and metamemory for name-face pairs. First, we can conclude that name-face pairs are governed by the same two memory processes (conscious and automatic) that have been observed in previous research. Secondly, timing in respect to the presentation of prospective memory judgments had been shown to make a significant contribution to the accuracy of metamemory predictions for names and faces, specifically when the judgments are presented at a delay. These results reflect the findings of previous studies on the subject (Daniels et al., 2008; Toth et al., 2011) and provide new insight into the nature of memory for names and faces. Overall, this research indicates that automaticity and timing of the judgments are underlying causes of memory inaccuracy for name-face pairs and provide new implications for future theory and memory research.

REFERENCES

Arbuckle, T., & Cuddy, L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, *81*, 126–131.

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In Spence, K.W. & Spence, J.T, Eds. *The psychology of learning and motivation (Volume 2)*. New York: Academic Press. pp. 89–195.

Begg, I., Duft, S., Lalonde, P., & Melnick, R. (1989). Memory predictions are based on ease of processing. *Journal Of Memory And Language*, *28*(5), 610-632. doi:10.1016/0749-596X(89)90016-8

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal Of Experimental Psychology: General*, *127*(1), 55-68. doi:10.1037/0096-3445.127.1.55

Benjamin, A. S., & Diaz, M. (2008). Measurement of relative metamnemonic accuracy. In J. Dunlosky, R. A. Bjork, J. Dunlosky, R. A. Bjork (Eds.) , *Handbook of metamemory and memory* (pp. 73-94). New York: Psychology Press.

Bowne, B. P. (1886). *Introduction to psychological theory*. New York: American.

Brown, R. & McNeill, D. (1966). The “tip of the tongue” phenomenon. *Journal of Learning and Verbal Behavior*, *5*, 325-377.

Bruce, V. (2009). Remembering faces. In J. R. Brockmole, J. R. Brockmole (Eds.) , *The visual world in memory* (pp. 66-88). New York, NY US: Psychology Press.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal Of Psychology*, *77*(3), 305-327. doi:10.1111/j.2044-8295.1986.tb02199.x

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review, 7,* 26–48. doi:10.3758/BF03210724

Carroll, M., & Buss, R. (1988). Fame attributions and the feeling of knowing. *Australian Journal of Psychology, 40*(1), 35-43. doi:10.1080/00049538808259067

Carroll, M., & Simington, A. (1986). The effects of degree of learning, meaning, and individual differences on the feeling-of-knowing. *Acta Psychologica*, *61*(1), 3-16. doi:10.1016/0001-6918(86)90018-1

Cavanaugh, J. C., & Perlmutter, M. (1982). Metamemory: A critical examination. *Child Development*, *53*(1), 11-28. doi:10.2307/1129635

Conant, R. C., & Ashby, W. R. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science, 1*, 89-97.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, *2,* 6. 671–684. doi:10.1016/S0022-5371(72)80001-X.

Daniels, K., Toth, J., & Hertzog, C. (2009). Aging and recollection in the accuracy of judgments of learning. *Psychology and Aging, 24*(2). 494-500. doi: 10.1037/a0015269

Do Lam, A. A., Axmacher, N., Fell, J., Staresina, B. P., Gauggel, S., Wagner, T., Olligs, J., & Weis, S. (2012). Monitoring the mind: The neurocognitive correlates of metamemory. *Plos ONE*, *7*(1), doi:10.1371/journal.pone.0030009

Dunlosky, J., & Nelson, T. O. (1992). The importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20,* 374–380. doi:10.3758/BF03210921

Gallo, D.A., Meadow, N.G., Johnson, E.L., & Foster, K.T. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Memory and Language, 58* , 1095-1111.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*(4), 309-313.

Gardiner, J. M. (2001). Episodic memory and autonoetic consciousness: A first-person account. Philosophical Transactions of the Royal Society London, 356, 1351–1361. doi:10.1098/rstb.2001.0955.

Godden, D., & Baddeley, A. (1980). When does context influence recognition memory? *British Journal of Psychology*, *71*, 99-104.

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, *49,* 732-764.

Greene, J. W., & Hodges, J. R. (1996). Identification of famous faces and famous names in early Alzheimer's disease: Relationship to anterograde episodic and general semantic memory. *Brain: A Journal of Neurology*, *119*, 111-128. doi:10.1093/brain/119.1.111

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*, *4*, 208-216. doi:10.1037/h0022263

Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior*, *6*, 685–691. doi:10.1016/S0022-5371(67)80072-0

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29,* 22– 34. doi:10.1037/0278-7393.29.1.22

Hockley, W. E., Hemsworth, D. H., & Consoli, A. (1999). Shades of the mirror effect: Recognition of faces with and without sunglasses. *Memory & Cognition*, *27*(1), 128-138.

Hunt, R., & Rawson, K. A. (2011). Knowledge affords distinctive processing in memory. *Journal Of Memory And Language*, *65*(4), 390-405. doi:10.1016/j.jml.2011.07.001

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513-541. doi:10.1016/0749-596X(91)90025-F

Jacoby, L. L., & Craik, F. I. M. (1979). Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S.Cermak & F. I. M.Craik (Eds.), *Levels of processing in human memory* (pp. 1-21). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Jacoby, L. L., Yonelinas, A. P., & Jennings, J. M. (1997). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. D. Cohen, J. W. Schooler, J. D. Cohen, J. W. Schooler (Eds.) , *Scientific approaches to consciousness* (pp. 13-47). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

James, L. E., Fogler, K. A., & Tauber, S. K. (2008). Recognition memory measures yield disproportionate effects of aging on learning face-name associations. *Psychology and Aging*, *23*(3), 657-664. doi:10.1037/a0013008

James, W. (1890). *The principles of psychology*. New York: Holt.

Jang, Y., & Nelson, T. O. (2005). How many dimensions underlie judgments of learning and recall? Evidence from state–trace methodology. *Journal of Experimental Psychology: General*, *134*, 308–326. doi:10.1037/0096-3445.134.3.308

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302-4311.

Kanwisher, N., & Yovel, G. (2006). The Fusiform Face Area: A Cortical Region Specialized for the Perception of Faces. *Philosophical Transactions: Biological Sciences, 361*, 1476. 2109-2128.

Kausler, D. H., & Lair, C. V. (1965). R-S (‘‘backward’’) paired-associate learning in elderly subjects. *Gerontology, 20*, 29–31.

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. doi:10.1037/0096- 3445.126.4.349

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(2), 187-194. doi:10.1037/0278-7393.31.2.187

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, *52*(4), 478-492. doi:10.1016/j.jml.2005.01.001

Lobmaier, J. S., & Mast, F. W. (2008). Face imagery is based on featural representations. *Experimental Psychology*, *55*(1), 47-53. doi:10.1027/1618-3169.55.1.47

Lockhart, R. S. (2002). Levels of processing, transfer-appropriate processing, and the concept of robust encoding. *Memory*, *10*, 5/6. 397-403. doi:10.1080/09658210244000225

Loftus, G. R., Oberg, M. A., & Dillon, A. M. (2004). Linear theory, dimensional theory, and the face-inversion effect. *Psychological Review, 111,* 835–863. doi: 10.1037/0033-295X.111.4.835

MacKenzie, G., & Donaldson, D. I. (2009). Examining the neural basis of episodic memory:

ERP evidence that faces are recollected differently from names. *Neuropsychologia*, *47*(13). 2756-2765. doi:10.1016/j.neuropsychologia.2009.05.025

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252-271. doi:10.1037/0033-295X.87.3.252

Masson, M. E. J., & Rotello, C. M. (2009). Sources of bias in the Goodman-Kruskal gamma coefficient measure of association: Implications for studies of metacognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 509–527. doi: 10.1037/a0014876

Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs?. *Memory & Cognition*, *29*(2), 222-233.

McCabe, D. P., Roediger, H. L., McDaniel, M. A., & Balota, D. A. (2009). Aging decreases veridical remembering but increases false remembering: Neuropsychological test correlates of remember/know judgments. *Neuropsychologia, 47,* 2164 –2173. doi:10.1016/j.neuropsychologia.2008.11.025

McCabe, D. & Soderstrom, N. (2011). Recollection-based prospective metamemory judgments are more accurate than those based on confidence: Judgments of remembering and knowing (JORKs). *Journal of Experimental Psychology: General.* Advance online publication. doi: 10.1037/a0024014

McWeeny, K. H., Young, A. W., Hay. D. C., & Ellis, A. W. (1987). Putting names to faces. *British Journal of Psychology, 78,* 143–149.

Metcalfe, J. (2000). Metamemory: Theory and data. In E. Tulving, F. M. Craik, E. Tulving, F.

M. Craik (Eds.) , *The Oxford handbook of memory* (pp. 197-211). New York, NY US: Oxford University Press.

Minear, M. & Park, D.C.(2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers, 36*, 630-633.

Naveh-Benjamin, M., Guez, J., Kilb, A., & Reedy, S. (2004). The associative memory deficit of older adults: Further support using face-name associations. *Psychology and Aging*, *19*(3), 541-546. doi:10.1037/0882-7974.19.3.541

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95*, 109–133. doi:10.1037/0033-2909.95.1.109

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science, 2*, 267–270. doi: 10.1111/j.1467-9280.1991.tb00147.x

Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science, 5*(4), 207-213. doi: 10.1111/j.1467-9280.1994.tb00502.x.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125-173). San Diego, CA: Academic Press.

Old, S. R., & Naveh-Benjamin, M. (2011). Age differences in memory for names: The effect of prelearned semantic associations. *Psychology and Aging*. Advance online publication. doi:10.1037/a0025194

Ornstein, P. A. , & Corsale, K.(1979). Organizabonal factors in chddren's memory In C. R. Puff (Ed.), *Memory organization and structure.* New York Academic Press.

Parkin, A. J., Gardiner, J. M., & Rosser, R. (1995). Functional aspects of recollective experience in face recognition. *Consciousness and Cognition: An International Journal*, *4*(4), 387-398. doi:10.1006/ccog.1995.1046

Parks, C. M., & Yonelinas, A. P. (2007). Moving beyond pure signal detection models: Comment on Wixted (2007). *Psychological Bulletin, 114*, 188 –202.

Pickering, E. C., & Schweinberger, S. R. (2003). N200, N250r, and N400 Event-Related Brain Potentials Reveal Three Loci of Repetition Priming for Familiar Names. *Journal of Experimental Psychology: Learning, Memory, And Cognition*, *29*(6). 1298-1311. doi:10.1037/0278-7393.29.6.1298

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 435-451. doi:10.1037/0278-7393.18.3.435

Rentz, D. M., Amariglio, R. E., Becker, J., Frey, M., Olson, L. E., Frishe, K., & ... Sperling, R. (2011). Face-name associative memory performance is related to amyloid burden in normal elderly. *Neuropsychologia*, *49*(9), 2776-2783. doi:10.1016/j.neuropsychologia.2011.06.006

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General, 137,* 615–625. doi:10.1037/ a0013684

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin*, *137*(1), 131-148. doi:10.1037/a0021705

Schmidt, I. W., Dijkstra, H. T., Berg, I. J., & Deelman, B.G. (1999). Memory training for remembering names in older adults. *Clinical Gerontologist, 20*, 57–73.

Sommer, W. W., Heinz, A. A., Leuthold, H. H., & Matt, J. J. (1995). Metamemory, distinctiveness, and event-related potentials in recognition memory for faces. *Memory & Cognition, 23*(1), 1-11.

Spellman, B. A., & Bjork, R. A. (1992). When predictions create reality: Judgments of learning may alter what they are intended to assess. *Psychological Science, 5*, 315–316. doi:10.1111/j.1467-9280.1992.tb00680.x

Tauber, S. K., & Rhodes, M. G. (2010). Metacognitive errors contribute to the difficulty in remembering proper names.*Memory*, *18*(5), 522-532. doi:10.1080/09658211.2010.481818

Toth, J. P. (1996). Conceptual automaticity in recognition memory: Levels-of-processing effects on familiarity. *Canadian Journal of Experimental Psychology*, *50*(1), 123-138. doi:10.1037/1196-1961.50.1.123

Toth, J. P., Daniels, K. A., & Solinger, L. A. (2011). What You Know Can Hurt You: Effects of Age and Prior Knowledge on the Accuracy of Judgments of Learning. *Psychology and Aging*. Advance online publication. doi: 10.1037/a0023379

Troyer, A. K., Häfliger, A., Cadieux, M. J., & Craik, F. M. (2006). Name and face learning in older adults: Effects of level of processing, self-generation, and intention to learn. *The Journals of Gerontology: Series B: Psychological Sciences And Social Sciences*, *61*(2), 67-74.

Tse, C., Balota, D. A., & Roediger, H. L. (2010). The benefits and costs of repeated testing on the learning of face–name pairs in healthy older adults. *Psychology and Aging, 25*(4), 833-845. doi:10.1037/a0019933

US Social Security Administration. (2011). *Popular baby names* [Online database]. Retrieved from http://www.socialsecurity.gov/OACT/babynames/.

Watier, N. N., & Collin, C. A. (2011). Metamemory for faces, names, and common nouns. *Acta Psychologica*, *138*(1). 143-154. doi:10.1016/j.actpsy.2011.05.017

Weaver, C., & Kelemen, W. (1997). Judgments of learning at delays: Shifts in response patterns or increased metamemory accuracy? *Psychological Science, 8*, 318 –321. doi:10.1111/j.1467-9280.1997 .tb00445.x

Weaver, C., & Kelemen, W. (2003). Processing similarity does not improve metamemory: Evidence against Transfer-Appropriate Monitoring *Journal of Experimental Psychology: Learning, Memory and Cognition, 29,*1058-1065.

Werheid, K., & Clare, L. (2007). Are faces special in Alzheimer's disease? Cognitive conceptualisation, neural correlates, and diagnostic relevance of impaired memory for faces and names. *Cortex:, 43*(7). 898-906. doi:10.1016/S0010-9452(08)70689-0

Wiese, H. (2011). The structure of semantic person memory: Evidence from semantic priming in person recognition. *British Journal of Psychology, 102*, 899-914. doi:10.1111/j.2044-8295.2011.02042.x

Wiese, H., & Schweinberger, S. R. (2008). Event-related potentials indicate different processes to mediate categorical and associative priming in person recognition. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*, *34*(5). 1246-1263. doi:10.1037/a0012937

Yesavage, J. A., Rose, T. L., & Bower, G. H. (1983). Interactive imagery and affective judgments improve face-name learning in the elderly. *Journal of Gerontology, 38*, 197–203.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal Of Memory And Language*, *46*(3), 441-517. doi:10.1006/jmla.2002.2864