THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF PSYCHOLOGY

RECOGNITION MEMORY FOR IDENTICAL, MORPHED, AND NEW FACES

ETHAN ZIMMERMAN
Fall 2011

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Psychology
with honors in Psychology

Reviewed and approved* by the following:

Nancy A. Dennis
Assistant Professor of Psychology
Thesis Supervisor

Richard A. Carlson
Professor of Psychology
Honors Advisor

* Signatures are on file in the Schreyer Honors College.

Abstract

Are eyewitnesses really reliable sources who should be used to identify criminals? That question has been highly debated because eyewitness testimony is still being used today as evidence in the conviction of criminals. It is therefore vital that these witnesses identify the correct criminal, but is that always the case? To understand the processes involved in episodic memory, it is pertinent to first realize that people encode significant amounts of information every day, whether they want to or not. Therefore, it is how this information is processed, sorted, and stored that allows people to selectively retrieve knowledge as necessary, which comes into play when witnesses must recollect their memory of a crime. Consequently, this study, in the context of false memory, aims to investigate that highly debated question and determine the degree to which people confused similar faces for those actually studied in a recognition memory task. In this study, participants viewed photographs of faces at encoding, and then saw target faces, related morphed faces at various gradients, or unrelated new faces at retrieval. Results showed that participants could not distinguish between a target face and a very similar morphed face. These results suggest that eyewitnesses are not well-able to distinguish between a suspect of a crime and someone who is extremely similarly looking to that suspect. Accordingly, it is theorized that memory recognition is in large part explained by the fuzzy-trace theory, as well as the dual process theory, as it is clear that memories are influenced by reconstructive processes. Future investigation of this issue will analyze recognition as a function of time delay.

ACKNOWLEDGMENTS

I am extremely grateful and honored to have been given the opportunity to work with my thesis supervisor and undergraduate research professor, Dr. Nancy Dennis, for the past four years. Her guidance, humor, creativity, and patience have allowed me to explore my interests and understand the subject matter of research in cognitive psychology, which has resulted in the completion of this honors thesis.

Also, I would like to thank Simon Vandekar for his assistance and support throughout this process, as his willingness to help others and his passion for statistics never went unnoticed during the completion of the project. He was always willing to help out regardless of what he was doing at the time.

TABLE OF CONTENTS

Introduction

The brain has the remarkable capability of facial recognition which proves important in everyday life for a variety of reasons (Van Belle, De Graef, Verfaillie, Busigny, & Rossion, 2010). It is believed that the dual process theory and the fuzzy-trace theory of memory play a significant role in the facial recognition process due to the complex nature of cognitive processes involved in such a task (Barrett, Tugade, & Engle, 2004). Clearly then, facial recognition requires both automatic and controlled processing, which is why even though humans possess an extraordinary ability to store a multitude of information in both short and long term memory, some of these memories, especially those regarding real life experiences, are prone to confusion and error. Therefore, like most information stored in memory, facial recognition and people's memory of faces are highly susceptible to the eventual compilation of false memories of faces because of the dual automatic and controlled memory processes as well as the formation of gist memory traces. Ultimately, it is important to understand facial recognition and its potential flaws if we are to understand its limitations of use in everyday life, especially if we are utilizing the testimony of eyewitnesses in the conviction of criminals. Specifically, there are explicit flaws involving police interviews of eyewitnesses, as eyewitnesses can be subjected to multitudes of misinformation (misinformation effect) where individuals are provided with information and details about an event or persons that are not completely accurate or truthful, but are eventually believed to be true by witnesses. It has also been found on numerous occasions that police interviewers do not "accept" a witnesses account and push the witness to provide a more "accurate" picture of what had occurred based on what investigators believe to be true (Brainerd & Reyna, 2005). Moreover, we can look at a more implicit explanation for such eyewitness flaws, which can be found relating to the misinformation paradigm of memory, in which individuals might witness an event, and an interpolated event might interfere with their original encoded event, thus creating a false memory and ultimately an inaccurate recognition of target faces (Brainerd & Reyna 2005). For example, it has been found that during police interviews, eyewitnesses are often subjected to memory suggestion, whereby false memories are implanted by officers' suggestions of what had actually occurred based on internal evidence found and/or based on other witnesses' testimony. Besides the dual process and fuzzy-trace theory, other mechanisms thought to underlie facial recognition are still poorly understood, and although this study does not aim to explicitly determine the mechanism behind facial recognition, there are some other mechanisms that have been proposed that specifically deal with false memories of composite faces. Composite faces, such as those used in the study are faces created by combining facial features of two different faces. These proposed mechanisms will be discussed later. For the most part however, facial recognition is one of many facets that make humans unique, and unfortunately, the brain's capacity for memory is prone to implicit reconstruction of its once original events. Such reconstructive processes can and will result in the formation of false memories. Therefore, an important question to consider is whether or not we should be using eyewitnesses' testimony as wholehearted evidence or even partial evidence when convicting suspects of crimes in a society today, when research points to the presence of errors in memory (Busey & Tunnicliff, 1999).

Consequently, it is important to discuss the practical implications of eyewitness testimony and the possible reasons why eyewitness testimony is not as accurate and reliable as people would like to think it is. After all, it is a difficult concept to believe that someone really would testify that a specific suspect was indeed the person who committed a crime in question if that witness was not sure of his or her decision. Additionally, it is equally disturbing to believe that witnesses who are very confident in identifying a suspect of a crime may not be correct in their decision either. Remarkably, both of these cases happen more often than people know about, and it is the sad truth that people are being sentenced to prison terms for crimes they never committed because an “eyewitness” picked them out through photograph, through actual person visualization, or by some other means. This occurs both when viewing individuals one at a time, or when viewing them in the more traditional route, using suspect line-ups, which has proven to be even less accurate than single viewing of suspects (Lindsay & Wells, 1985). Ultimately, it has been found that the primary reason for the wrongful conviction of innocent people is based on misidentification from a lineup or photospreads (Wells & Bradfield, 1998). Additionally, it has been estimated that the overall wrongful conviction rate is between 0.5% and 5.0% (Huff, Rattner, & Sagarin, 1996), and approximately 75% of the wrongful convictions are in some way related to eyewitness misidentification (McGrath, 2011). Therefore, even though the percentages of wrongful conviction based on eyewitness testimony is not exactly known, it is clear that suspect misidentification does occur fairly frequently and it should seem unacceptable to have even one innocent person sent away to prison for a crime he or she did not commit. One specific instance where a man was sentenced to two consecutive life terms in prison because of eyewitness misidentification occurred in Texas, where Thomas McGowan was “chosen” out of a lineup as a victim’s rapist. McGowan was subsequently sentenced to prison where he spent 23 years there as of 2008. In 2008, he was finally released after DNA evidence exonerated him. In this case, the victim who chose McGowan from the lineup felt pressured by the police to provide an answer when shown a list of photographs of possible suspects. She was also not very confident in her decision, yet was told that she had to be confident if they were going to convict the man and use her testimony as evidence. So unfortunately, the misinformation and outside influence played a part in this man losing essentially his entire adult life, something that absolutely needs to be avoided in the future (The Innocence Project, 2008). Fortunately, with new technology today, those wrongful conviction rates due to misidentification are decreasing based on analysis of DNA and other physical evidence. Nevertheless, the problem still remains, and because there are in fact accurate and valuable eyewitnesses, it is hard to determine a direction to go with eyewitness testimony. But first, we must consider from a psychological standpoint why people identify the wrong person when asked to identify the person or persons involved in a crime. Unfortunately, there is no one answer nor is there one simple answer. Subsequently, although a witness may develop a memory trace of a person or event, research has shown that people have poor memory for new or unfamiliar faces, which is typically the type of face a witness will encounter during the time of a crime (Hancock, Bruce, & Burton, 2000). Additionally, just like traces of physical evidence, such as DNA, fingerprints, or blood stains, memory traces of an event or person are subject to contamination, destruction, and change (Wells & Loftus, 2003). Because we have no video recorder programmed into our brain or our memory, it is next to impossible to remember all details specific to an event or a person. Therefore, it is clear that the memory traces are susceptible to reconstruction.

In fact, memory traces often times result in false recollection, which occurs when people recall information predominately based on familiarity of an item (faces), and less on specific perceptual details of the items (faces). As a result, people will falsely recollect information, but to them, the information they recollect feels like an actual portrayal of the information they once encoded. Likewise, people will even recollect false information with high confidence and at high response rates (Dennis, Bowman, Vandekar, 2011).

These reconstruction processes might be a result of the misinformation effect (referenced above), where events such as input from police officers about a crime, information published in a newspaper or online, the result of information overload in one's short term memory, the specific memory simply disappears over time, as do most memories, or a wide array of other possibilities influence one's memory. For example, in Loftus' eyewitness study of an automobile accident, participants were asked questions about the accident, but deliberate wording of the questions caused participants to recall information that never occurred (Loftus, 1993). Overall, it is clear and evident that although the human brain has a huge capacity for memory, memories are not always stored as exact replicas of what was actually seen and are actually widely susceptible to reconstructive processes.

The current study aims to investigate the degree to which people confused similar faces for those faces actually studied. Our primary interests focused on morphed faces used in the study and the facial recognition errors that resulted from trying to recognize original faces from the similar-looking morphed faces.

Typical memory studies employ novel distracters, a type of distractor made up of only parts of non-studied items. More specifically to this study, a novel distractor is a face that was constructed and used from only non-studied faces during the old/new recognition test. Accordingly, research shows that when novel distractors are used facial recognition errors result (Busey & Tunnicliff, 1999). However, this study primarily utilizes feature distractors instead of novel distractors, which means that faces seen during the old/new recognition task are constructed by combining information and features of one studied face and one non-studied face. In this study, the morphed faces are exactly that, feature distractors in which all the morphed faces were creating by combining features from two different faces; one of which was studied during encoding, and one of which was never shown. Therefore, facial recognition errors are likely to occur because the morphed feature distractors are even more similar to the previously studied items. One hypothesis for these recognition errors has been mainly focused on conjunction errors, in which participants use a specific mechanism to recombine studied items, which in turn, creates novel memory traces. As a result, the novel traces are then perceived as the conjunction stimuli, stimuli created from features of two different studied items, during an old/new recognition test. Ultimately, this misperception would result in false alarms. It is proposed that there is a binding error that results from the failure of an inhibitory mechanism that would normally prevent novel traces from just formed sporadically (Kroll, Knight, Metcalfe, Wolf, & Tulving, 1996). However, because this study utilizes feature distractors in the morphed stimuli instead of conjunction distractors, false alarms may not necessarily be a result of the same mechanism and subsequent mechanism/binding error. Therefore, it might be possible that recombination does indeed occur during the study phase; however, the inadvertently recombined memory traces of faces do not interact the same way during the recognition test. Thus, the recombination that might take place during encoding would not necessary match up with the feature stimuli at test, enabling participants to hypothetically make slightly more accurate judgments.

Furthermore, there is another theory related to memory in which retrieval items are composites of two different items. For one, facial recognition could also be a basis of what is called the familiarity process, an element of the dual process theory (Verde, 2010). Based on this process, participants see the retrieval stimuli (i.e. the morphed faces) and they look for retrieval cues that might relate to the contents of short term memory. Therefore, when participants see the feature distractors, it is possible that participants see the images and search for some context that is familiar to an item in their short term memory, whether it be the faces shown at encoding or other faces already stored in memory. As a result of the apparent similarity of the morphed faces to the studied faces, this familiarity process would likely contribute to high rates of false alarms to the morphed faces.

Additionally, along with the theory of familiarity process, the recollection process comes into play. This process is more discriminant and slower and focuses on searching for specific details of the episodic memory. Therefore, the recollection process works to help the participant distinguish certain features of the retrieval items from the studied items; however, if the features of the morphed faces are very similar to the studied items, it may be difficult for one to perceive any feature differences and thus there would be an increasing number of false alarms as well. The familiarity process has been shown to be quicker than the recollection process, and is considered to be more of an implicit process. Alternatively, recollection is more similar to free recall, and is less sensitive to time interval delays.

Moreover, participants are not expected in such a short amount of time and without prompt to remember the exact details and features of an old face presented at encoding, as would be the case with randomly witnessing a criminal event. Rather, they are predicted to encode the entire face as a whole, what is known as a holistic view of recognition. This view emphasizes the basis of facial recognition that specific facial features are not perceived and represented independently of one another, but is instead, integrated as a face at whole, without regard to specific feature modifications (Van Belle et al., 2010). This is a result of memory errors that occur from lack of efficient processing at encoding, or simply the lack of encoding all together. Therefore, without an adequate and proficient encoding of individual facial features, individuals would perceive similarly-looking morphed faces during retrieval as old faces.

Ultimately, it is evident that there are a few different concepts that can be examined through this false memory study. For one, based on the notion and prevalence of conjecture errors in a mixed-stimuli design as this, it is predicted that participants will have a high false alarm rate of facial recognition for faces that were morphed and presented during retrieval. Ultimately, due to the apparent similarity of the morphed faces, even as they contain less and less similar character of the parent faces of which they were morphed from, participants should perceive the morphed faces as old faces. Additionally, participants are not expecting a memory test after encoding, so all encoding that occurs is incidental from the encoding task. Furthermore, we expect, and would hope, that participants have a better recognition memory for all faces presented at retrieval when they rate that they are highly confident in their decision. This is because it takes two steps, one to decide if a face is old or not and then one step to determine confidence. If the participant decides that a face was shown during encoding and is highly confident, it is expected that these individuals, or any individuals for that matter, would be more accurate in his or her decision. After all, eyewitnesses and their testimony are more accurately utilized in criminal cases when participants are very confident in their decisions. In fact, as we saw in the McGowan case, witnesses are told that they have to be confident. Therefore, we predict that participants making highly confident decisions would be able to distinguish faces better than those participants who might have guessed or participants who might not have been quite sure in their memory judgments. If that were the case, and confidence did not make a difference based on facial recognition, one could argue that eyewitnesses should no longer be used a means of conviction.

Methods

Participants. Twenty healthy undergraduates [sixteen female; mean age = 18.9(1.02)] attending The Pennsylvania State University were recruited from The Pennsylvania State

University’s (PSU) psychology subject pool and received course credit for participation.

Stimuli. The stimuli were 126 pictures of young adult male faces with neutral expressions. Seventy-six of the men had facial hair. Half of the faces were Caucasian and the other half were a mix of minority races, including African American, Indian, Asian, and Hispanic. 120 faces were used to create stimuli for the study portion of the experiment, and the other 6 faces were only used during the retrieval portion of the study as unrelated lures. At retrieval, 60 of the original 120 faces were used as parent faces (parent X faces), original faces that were combined at varying percentages with the other 60 parent faces (parent Y faces) to create 240 related morphed faces in total (4 morphed faces per parent pair). The parent faces were simply original faces used for the morphing process and were faces that were presented as targets during retrieval, as they remained unaltered. All faces were morphed within the same racial grouping and the faces were blended at percentages of (80:20), (60:40), (40:60) and (20:80), with the first number in each set representing the percentage of one parent face in a morphed face and the second number in each set representing the percentage of the second parent face in the morphed face. 7 of each morph blend, for a total of 28 were used during retrieval as related lures. Also, 32 parent faces from each set (either Parent X or Parent Y) were used as targets during retrieval. See figure 1. The faces were counterbalanced as to which faces participants would see. The 32 faces in one counterbalance were morphed with the set of 32 other parent faces in the second counterbalance. More specifically, parent face 1 of counterbalance 1 was morphed with Parent face 1 of counterbalance 2. Therefore, in counterbalance 1, Parent face 1 is shown, and the subsequent morphed faces; however, Parent face 1 of counterbalance 2 is not shown. The reverse is true for counterbalance 2. See figure 1 for clarification.

The morphed stimuli were created by categorizing images obtained from The Color FERET

Database and The AR Face Database by race and selecting two faces of the same race to be morphed. Face pairs were selected semi-randomly, as it was not possible to fully randomize the selections because of different races and different facial hair features. The faces were morphed together using Fantamorph, a photo morphing software made by Abrosoft. All images were cropped to a size of 3.958 inches by 5.313 inches and were placed on a white background using Adobe Photoshop Elements.

Design and Procedure. All participants were given a consent form to sign that was approved by The Pennsylvania State University's Institutional Review Board (IRB) and participants were informed that the study would last approximately thirty minutes. Stimuli were presented on a computer screen and participants made responses using the keyboard. There was a ten minute distractor period in between encoding and retrieval.

The study phase was counterbalanced for which parent faces were used and subsequently shown to participants. Each counterbalance contained half Caucasian and half minority faces. Subsequently, during the study phase of the experiment, participants viewed 60 images of male faces presented one at a time. Each face was presented for 10 seconds, whereby the participants were instructed to think about how typical each face was and then make a 5 point "typicality" rating following the presentation of each face. Participants were instructed to make a rating of a "1" to indicate that the face was very atypical, and therefore distinguishable and easy to pick out of a crowd. Participants were also told to use a rating of "5" to indicate that the face was highly typical, and thus difficult to pick out of a crowd. Participants were encouraged to use all 5 keys (1-5) in making their rating and had no time limit on making their typicality decision. The study phase consisted of two blocks, each consisting of 30 faces. The entire study phase lasted approximately twenty minutes.

During the ten minute distracter period following the study phase, participants completed a demographic questionnaire and the WAIS Matrix Reasoning Test (Wechsler Adult Intelligence Scale) third edition (see Wechsler, 1997). After the distracter period, participants were instructed about retrieval.

The retrieval phase of the study contained 66 total faces. Each counterbalance contained 32 original parent faces, 28 related morphed faces, which were divided equally into seven (80:20) faces, seven (60:40) faces, seven (40:60) faces, and seven (20:80) faces, and six unrelated lure faces. (see Figure 1). Therefore, during retrieval, each participant completed a memory test in which they had to decide whether the face that appeared in the second part of the study was a face that was exactly identical to a face presented during the first session (old), or whether the face has never been shown before (new). There was no time limit for the participant to make his/her memory judgment. The faces presented during retrieval consisted of six different trial types (0-5), as shown in Figure 1. After participants had made their old/new memory judgment, they were then asked to rate their confidence on a scale of 1 to 3, using the number pad keys. After retrieval, participants were asked to rate their overall attentiveness and effort using a 9 point scale. The attentiveness and effort data was not factored in during analysis.

Results

For collapsed confidence responses, the mean hit rate was 0.779 (11.19) for target faces and the false alarm rate was 0.672 (25.00), 0.521 (20.35), 0.343 (20.41), 0.171 (21.54) respectively for each of the four morphed face trial type distracters (1-4). The mean false alarm rate for non-morphed distractors was 0.100 (8.38), see table 1. Additionally, in the collapsed confidence condition, participants' responses indicated a significant linear increase in false response rate as a function of similarity (via morphing) as seen in Figure 2 ($p < .001$).

Furthermore, participants' responses were subject to a planned T-Test with a Bonferonni correction at $p < 0.05$ and there was no significant difference in response rate for hits compared to the false alarm rate for the 80% morphed face ($t(19) = 1.358$, $p = 1.000$) in the collapsed confidence (see table 1). Moreover, for the collapsed confidence condition there was no significant difference in false alarm rates between 80% morphed faces and 60% morphed faces ($t(19) = 3.341$, $p = 0.275$), between 60% morphed faces and 40% morphed faces ($t(19) = 3.208$, $p = 0.070$), and between 20% morphed faces and new non-morphed faces ($t(19) = 1.567$, $p = 1.00$). Consequently under collapsed confidence condition participants were less likely to distinguish between faces with 20% difference in morphed character, indicating that when faces share 20% character of another face, people have a hard time distinguishing those faces; however, it is evident that when shared character between any two faces reaches 40%, participants were able to distinguish with some confidence that there were differences in the faces presented.

Alternatively, for high confidence responses, the mean hit rate was 0.595 (16.05) and the false alarm rate was 0.529 (22.28), 0.286 (17.34), 0.121 (11.61), 0.050 (9.58) respectively for each of the four morphed face trial type distracters (1-4). The high confidence false alarm rate for non-morphed distracters was 0.050 (7.84), see table 2. Also, when only high confidence responses were considered, the differences between all morphed trial types indicated a significant quadratic pattern of false response rate as a function of similarity (via morphing) as seen in Figure 3 ($p < .001$). There was also no significant difference in response rate for hits compared to the false alarm rate for the 80% morphed face in the high confidence condition ($t(19) = 1.115$, $p = 0.237$). In the high confidence condition, participants were not able to distinguish between a parent face and an 80% morphed face, as previously stated, and between a 20% morphed face and a new non-morphed face ($t(19) = 1.584$, $p = 1.00$). Overall, participants were not able to significantly distinguish between an encoded face and a similar 80% shared character morphed face (collapsed confidence: $t(19) = 1.358$, $p = 1.000$; high confidence: $t(19) = 1.115$, $p = 0.237$). However, other than the 80% morphed face, participants were able to distinguish between said parent face and all other trial type.

Discussion

The main purpose of this experiment was to determine how well people were able to distinguish between target faces and closely associated distractor faces that were controlled via morphing. Because morphed faces were used at varying degrees in the study, it was possible to determine at which degree people were able to distinguish between an old face and a new face (related morph lure or new unrelated lure). At the same time, participants' confidence on their memory judgment was taken into account to examine if individuals were actually able to make better recognition judgments when they were more confident. The results of the experiment demonstrated that morphed items that are more similar to the studied parent faces can attract a very high number of false alarms and that closely related faces (80:20) are indistinguishable from target faces; however, as a whole, all other feature distracters (morphed faces other than the (80:20) faces) did not elicit more false alarms than when participants viewed target faces. In other words, the (80:20) morphed faces were the only lure face responded to as 'old' as often as target faces. It was only when related lures were (60:40) faces that participants responded 'old' at a rate significantly different than when participants responded to target faces. From another standpoint, it is still evident that people are indeed highly skilled at face memory, with a 78% and 60% hit rate for collapsed and high confidence conditions, respectively (see table 1 and 2). Also, it was interesting that results suggested that we essentially saw the same pattern of hit rates and false alarms rates for the (80:20) faces in both the collapsed and high confidence conditions, indicating that facial recognition errors were not a result of guessing. Moreover, for the morphs and parents, the results are slightly different between response types for both conditions. The results indicate that the more similar the morphed face is to the parent face, the higher the false alarm rate is. However, the differences between response types indicate a different trend in response patterns (see Figure 2 and Figure 3). However, the false recognition overall confirms the notion that there may been some underlying psychological conjunction mechanism (see introduction) that works to monitor and combine the different features of a studied item and then allows an individual to formulate a new mental representation of what the face might be when presented with similar distracters, thus possibly causing an individual to incorrectly match the distracter to a encoded face if this 'new' morphed face appears to be 'old' face due to implicit memory reconstruction (Reinitz, Morrissey, & Demb, 1994). Ultimately, this mechanism could then be considered responsible for such high false alarms to non-studied faces.

In terms of eyewitness memory, it is evident that a large portion of external validity is compromised by conducting a study of this nature in a laboratory setting without actual crimes being committed or even being simulated really. Nevertheless, it was evident that the feature distracters of the morphed faces elicited a large number of false alarms across the board. However, in both collapsed and high confidence conditions, participants were actually able to distinguish between a parent face and all morphed faces, except the most "nearest neighbor" face in the 80% morphed face. Nevertheless, 33% of the time overall, participants were inaccurate in their facial recognition judgment when both fairly confident and very confident responses were considered. Also, when considering all responses, 27% of participants failed to make a correct facial recognition when they reported that they were very confident in their memory decision. When applied to eyewitness testimony, results suggest there was a lot of 'eyewitness' misidentification across the board, with participants inaccurately making recognition judgments approximately a third of the time. Subsequently, the results of this study could theoretically extend to how accurate eyewitnesses perform in the ‘real world.’ Although no studies have been found (that we know of) that have determined quantitatively how physically similar two people are or can be, it is evident that participants in the study had a difficult time distinguishing between a parent face and a very similar face. Therefore, it might be possible to consider the 80% morphed face in the context of eyewitness testimony as a potential twin brother or close cousin of the suspect in question. With that said, individuals in this study were not able to distinguish between said twin brother or close cousin and the actual suspect. Consequently, when faces are extremely similar to one another, individuals cannot tell the difference. As a result, the wrong people might be ‘chosen’ as the person who committed a crime.

Ultimately, one contemporary theory regarding facial recognition that emphasizes memory being subjected to both automatic and controlled processes that results in recognition errors is the dual process theory. One aspect of the dual process theory provides that individuals may turn to implicit memory and automatic retrieval processes based on the familiarity or prior memory of a “similar” face. Familiarity of a face is defined as remembering (typically implicitly as mentioned before) that is associated with vague, non-specific, and general feelings that an item, in this case a face, was part of or entirely an aspect of remembering a target face (Brainerd & Reyna, 2005). Therefore, if this aspect of the dual process model is accurate, people were making memory judgments of target faces, morphed faces, and new faces based on mechanisms that triggered a familiarity reaction of the presented face. In addition to familiarity, this theory emphasizes the fact that participants would have most likely also used controlled processing to recall specific details of faces that would allow individual participants to recognize a face through what is known as recollection. During recollection, remembering of faces involves memories of specific, vivid elements that were encoded when the target item (a parent face) was initially encoded, therefore allowing a participant to recollect specific events of their experience when the target face was presented (Brainerd & Reyna 2005). Familiarity is believed to be a quicker process than recollection, as recollection allows one to comprehend that an item (face) was previously encountered (Wixted, 2007). Together, both processes may lead to participants making their final memory judgment. However, assuming both processes did occur simultaneously or one was more strengthened than the other, it is suggested that memory errors would likely result, especially if participants are using familiarity as a basis for judgments more so than recollection. Consequently, it has been proposed by Yonelinas that recollection is a "high-threshold process" meaning that recollection either occurs or does not occur and that there is no in between. Therefore, if recollection does occur, individuals are more likely to provide a memory judgment that results in a high confidence decision. Contrastingly, if recollection fails (as it is proposed to be an all or none action), then familiarity kicks in during the remembering process and a response would be likely made with lower confidence (Yonelinas, Kroll, Dobbins, Lazzara, & Knight, 1998). Therefore, results showed that when participants were presented with (80:20) morphed faces, participants regarded those faces as 'old' just as often as target faces for both the collapsed and high confidence conditions. In regard to the dual process theory, it would explain that because confidence did not play a significant role in the false memory rates between a target face and (80:20) face, that participants most likely relied on familiarity of the morphed faces when making their memory decisions. This is because the (80:20) face in particular was extremely similar to the originally presented target face and thus was more accessible than any other morphed or new face in the participants' memory during the retrieval memory task. All in all, if recollection did not occur at a high rate as was probably the case in this study, then as the morphed faces became less and less similar to the target face, it is obvious that individuals would have a less intense familiarity reaction, thus enabling participants to better distinguish between a target face and faces with 60% shared character or less.

Furthermore, we can also turn to the fuzzy-trace theory as a mechanism behind why false memories of faces occurred in this study. The fuzzy-trace theory points to the presence of both verbatim and gist traces. Verbatim traces being representations of item-specific information and are detailed, whereas gist traces are representations of semantic information, are less precise, and are more "fuzzy" in our memory.   Gist retrieval supports false memory formations and results in memories of "false-but-gist-consistent" items (related morphed lures in this study). Thus, after viewing target items during the encoding task, participants develop both traces; however, storage of verbatim traces occurs too rapidly after viewing a target face and results in a disintegrated trace. As a result, gist traces are predominately activated during a retrieval task. For instance, if participants viewed a list of words at encoding: diary, horse, milk, and barn, and then were shown the word cow at retrieval. Participants' gist traces might be supported by a non-specific memory of "farm words" and therefore, when cow is presented at retrieval, participants respond indicating that the word is an old word because cow fits in with the other words on a non-specific level. When applied to this study, participants might have developed a gist trace of the target faces presented at encoding, but when shown morphed faces at retrieval, such traces might be activated, thus alerting the participant that such faces are old. However, it is evident that gist traces did not affect responses when viewing (60:40), (40:60). (20:80), and new faces, as participants could accurately distinguish between those faces and target faces. Subsequently, the gist trace formed at encoding caused individuals to see (80:20) faces as old because the gist of the target face was applied to the very similar face (Brainerd & Reyna, 2005).

Consequently, if witnesses rely primarily on familiarity reactions and gist traces, it is likely that memory for a face or event is without a doubt going to be a well-tainted reconstructed memory. As a result of the reconstruction, false memories will occur at high rates and with high confidence, which can lead to eyewitness misidentification. Therefore, because confidence did not play a significant role in people's facial recognition accuracy, it is evident that witness testimony should not be used solely as evidence in the conviction of criminals, as reconstructive processes along with outside interferences significantly compromise the accuracy of a sworn statement.

Caveats

As with all studies, there were limitations that may have affected the results and overall data collection. For one, the stimuli created using the Fantamorph morphing software did not allow for realistic morphing between females. This was because of the longer hairstyles of females, which did not allow for real human-looking stimuli. Subsequently, no female faces were used in the study. Therefore, it was possible that an own gender bias accounted for the high false alarm rates amongst the similar 80% morphed faces. The own gender bias provides a theory that individuals of one gender are better able to identify faces of the same gender (Wright & Sladden, 2003). Additionally, it has been found that encoding one's hair accounts for approximately half of the own gender bias when measured using hit and false alarm rates. Therefore, because the majority of participants (n = 16) were female, it may have been possible that the recognition rates were lower because no female faces were used. Subsequently, if a more representative sample was used, it might have been possible that recognition rates would have actually increased and there would have been no significant difference found in facial recognition between a parent face and an 80% morphed face.

Alternatively, recognition could have been skewed because only young male faces were used in the study as stimuli. This poses the question of own age bias, which alludes to the fact that individuals can recognize faces of their own age better than faces of another age (Anastasi & Rhodes, 2006). Therefore, because all participants were young college students and all the stimuli were of young males, recognition rates may have been higher because there were no older adults or children present. As a result, if the stimuli included males and females of all ages, the recognition rates might have decreased; however, this would not necessarily account for increased false alarm rates across the board.

In addition, significance of facial recognition could have been falsely high in this study because there were only a limited number of stimuli used during retrieval. Due to the small number of facial stimuli presented, it was much easier for slight differences in recognition rates to result in biased results. For instance, if one participant false alarmed to 4 out of 7 80% morphed faces, and also false alarmed to 3 out of 7 60% morphed faces, analysis of the data indicated that the individual was able to significantly distinguish between those two faces, when in fact, the individual only made one more correct response in the 60% condition as compared to the 80% condition.   Therefore, if more faces were used in the study, the results may have been more widespread, representative, and overall more accurate.

In addition, there may have been a time constraint that limited the study. There interference period for this particular study was only 10 minutes, which by no means represents any normal amount of time when a person might witness a crime and then be called in to make an attempt at identifying a suspect. Therefore, participants would obviously have better recognition memory for old faces they had just seen 10 minutes prior compared to faces that might not be seen again for hours, weeks, or even months.

Overall, there are some ways to that would help eliminate some of the limitations and biases to the study. For one, male and female faces should be included. In addition, older male and female faces should have been presented at encoding and retrieval. However, by adding in female faces, the realism of the stimuli and study would be compromised, which is something we tried to maintain by eliminating female faces. Nevertheless, by adding in female faces and older faces, the overall amount of stimuli would increase and therefore the different trial types would have significantly more faces in each type, thus reducing the sensitivity at which significant differences can be achieved.

Nonetheless, it is difficult to predict the validity of a study in this context as it is difficult to know how people will actually respond in real life situations when a crime has indeed been committed, as there are more contextual factors that affect one's memory storage and retrieval.

Future Directions

One interesting future direction for this study will be that it will be modified in order to examine the effects of facial recognition memory as a function of time delay. Therefore, the same experiment will be conducted; however, instead of a 10 minute interference delay between encoding and retrieval, there will be a one day delay experiment, as well as a one week experiment. By increasing interference time, we hope to extend the validity of the experiment to a more realistic eyewitness scenario. Thus, with an increase in time, participant's trace memory is expected to be subject to more contamination, destruction, and change as a result of a longer interference period, which is what would happen under 'normal' eyewitness scenarios. Therefore, we hypothesis that with an increase in time before the recognition period, participant's recognition memory for faces will decrease and they may actually be more likely to judge a more dissimilar face to the parent face as an old face. In other words, participants might now view 60% morphed faces or 40% morphed faces as identical to parent faces, etc.

Additionally, it would be interesting to examine the other-race effect, which states that individuals have a more difficult time recognizing faces of a different race (Sangrigoli, Pallier, Argenti, Ventureyra, & Schonen 2005). Although, there were equal numbers of Caucasian and other race faces, the participants were primarily Caucasian so it would be of interest to see how well these individuals actually did perform on their facial recognition task if Caucasian and other race faces were separated.

Furthermore, it would be important to examine the effects typicality had on recognition memory. As participants had to make a typicality rating during encoding based on how typical or common each face was, it would be interesting to examine whether or not the ratings and perceived typicality during encoding translating into better or worse facial recognition during retrieval when both studied items and new (morphed) items were presented. Previous studies (Bartlett, Hurry, & Thorley, 1984) have found that faces perceived as typical or common during the initial encoding actually elicit worse recognition memory when those same faces are presented at a later time during retrieval. However, a question arises then whether or not the typicality of a studied parent face has any effect on the recognition of a morphed face that was made from the same parent face when it is presented during retrieval. Therefore, it would be interesting to study whether or not a typical parent face translated into a typical morphed face, and thus this across stimuli typicality might have accounted for a decrease in hit rate, and a subsequent decrease in false alarm rate of the morphed faces if people . Alternatively, an atypical parent face might translate into an atypical morphed face and thus hit rate would increase, but false alarms would be less than expected, as participants are now viewing morphed faces as uncommon and thus are biased or are predisposed to select 'new' when presented with morphed faces. Because the morphed faces are perceived as uncommon, prior research gives evidence that facial recognition is more accurate when a face is 'distinct' in some way (Valentine & Ferrara, 1991). This would account for why individuals were fairly accurate in distinguishing between a studied parent face and all of the morphed faces, except the 80% face. In fact, the 80% face might have been so similar to the parent face, that a highly typical rated face during encoding translated equivalently to a highly typical morphed face at this degree. Therefore, participants would equally select a parent face and an 80% morphed face as an 'old' face. Ultimately, to determine the effects of typicality, the different ratings of each parent face would need to be analyzed in accordance with each parent face's set of morphed faces that were created.

Clearly, it is apparent that individuals in this study had a somewhat difficult time with facial recognition when the morphed faces were extremely similar to the studied parent face (the 80% condition). Nevertheless, it was evident that participants could indeed distinguish between the studied parent face and all the other morphed distracters and non-morphed distracters. Additionally, it was found that although there was an overall decrease in false alarms in the high confidence condition as compared to the collapsed confidence condition, there was no significant difference between facial recognition of studied parent faces and 80% morphed faces. Therefore, participants' confidence in their memory judgments did not have a significant effect in this study. Additionally, it was found across the board when all responses were considered that almost all participants had difficulty distinguishing between faces that had a 20% difference in shared character; however, recognition rates were much higher between trial types when faces were 40% or greater in terms of differing shared character. Ultimately, it is clear that participants were fairly accurate in their memory judgments overall, in that they were able to significantly distinguish between the parent faces and almost all of the morphed faces. But by far the most important entity learned through conducting this study is that there are many processes involved in the storage and recognition of faces in memory, which can ultimately lead to false memories, especially when faces are of particular similarity to a target face. It is with that knowledge, we should question the validity and strength of eyewitness testimony used in today's court system.

Trial Types	0	1	2	3	4	5
Mean Hit Rate	0.779	0.672	0.521	0.343	0.171	0.100

Table 1: Mean hit rate for the different trial types for collapsed confidence responses.

Trial Types	0	1	2	3	4	5
Mean Hit Rate	0.595	0.529	0.286	0.121	0.050	0.050

Table 2: Mean hit rate for the different trial types for high confidence responses.

							
% of parent faces	(100:0)	(80:20)	(60:40)	(40:60)	(20:80)	(0:100)	(0:0)
Stimuli Type	Parent X	Morph	Morph	Morph	Morph	Parent Y	NEW
Trial Type	0	1	2	3	4	--	5

Figure 1. Morphed faces created from combining two parent faces (Parent X and Y). The above stimuli represent examples of faces presented during the encoding and retrieval phases of the study for counterbalance 1. The only faces actually shown to participants in counterbalance 1 are trial types 0-5, as shown above. Alternatively, if we were to consider counterbalance 2, Parent Y would then become trial type 0, and the 20(80) face (shown as trial type 4) would be used as a 80(20) face (shown as trial type 1) and so on, excluding Parent X.

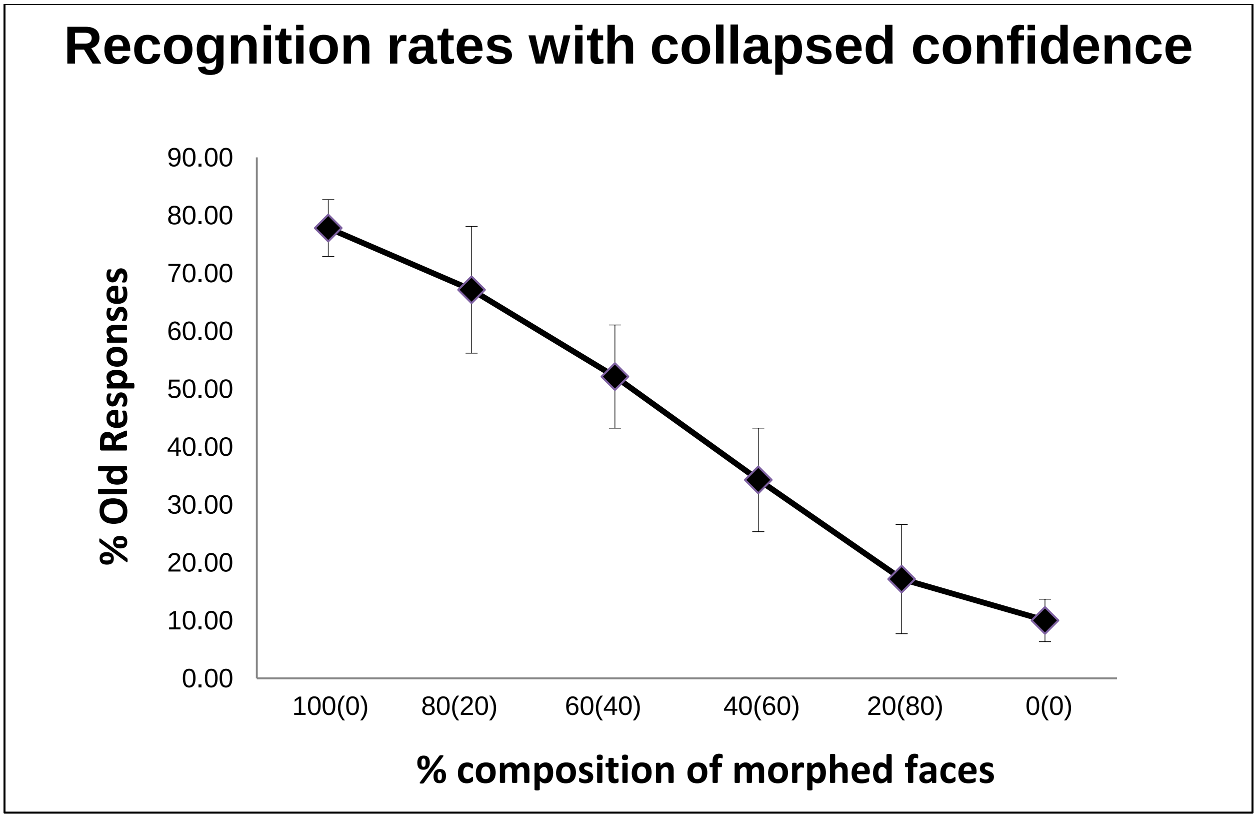

Figure 2. Percentage comparison of old responses based on different faces, including morphed faces, shown at retrieval. Responses of "fairly confident" and "very confident" decisions were included. Error bars represent 95% confidence intervals.

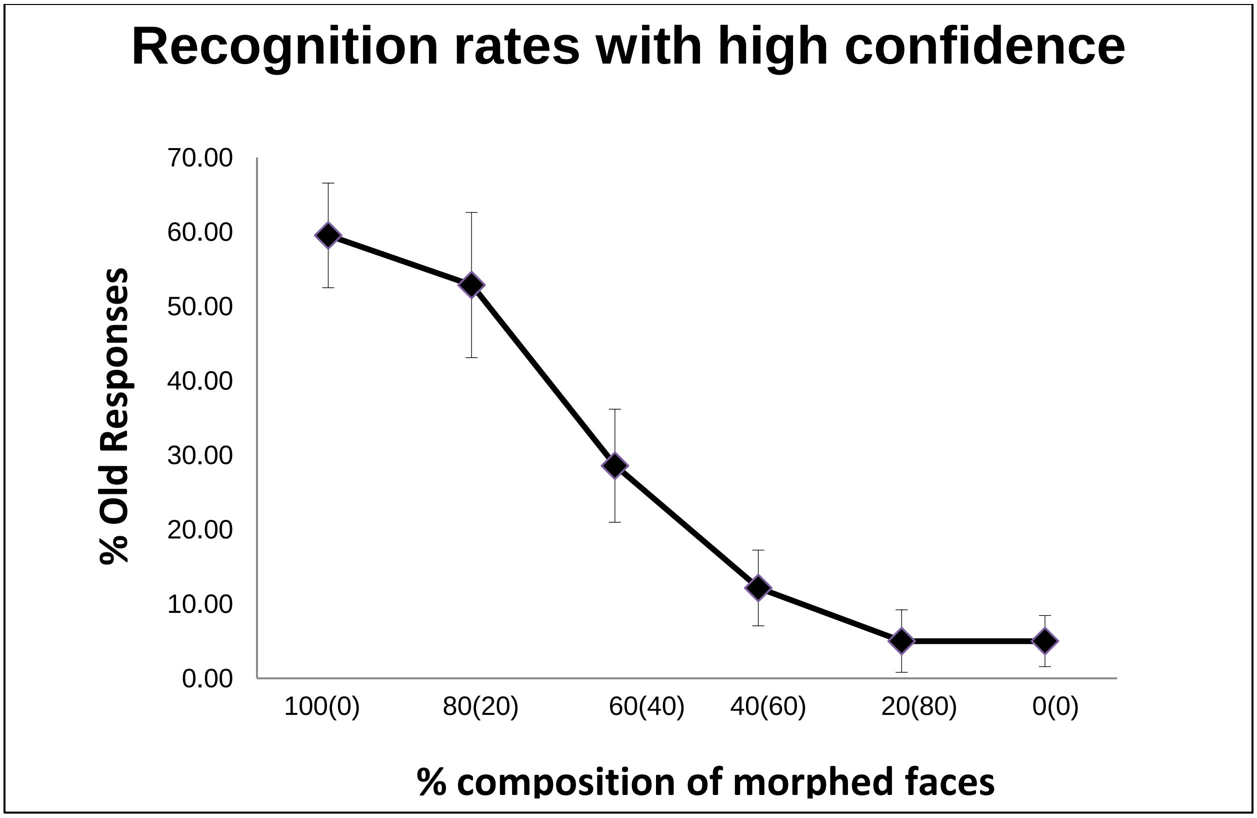

Figure 3. Percentage comparison of old responses based on different faces, including morphed faces, shown at retrieval. Responses for only "very confident" decisions were included. Errors bars represent 95% confidence intervals.

References

A. M. Martinez and R. Benavente. *The AR Face Database*. CVC Technical Report #24, June 1998.

Anastasi, J. S., Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review*, 12(6), 1043-1047.

Bartlett, J. C., Hurry, S., Thorley, W. (1984). Typicality and familiarity of faces. *Memory and Cognition*, 12(3), 219-228.

Barrett, L. F., Tugade, M. M. & Engle, R. W. (2004) Individual differences in working memory capacity and dual process theories of the mind. *Psychological Bulletin*, *130*, 553-573.

Bussey, T. A., Tunnicliff, J. L. (1999). Accounts of Blending, Distinctiveness, and Typicality in the False Recognition of Faces. *Journal of Experimental Psychology*, 25(5), 1210-1235.

Dennis, N.A., Bowman, C.R., & Vandekar, S.N. (In press). True and phantom recollection: An fMRI investigation of similar and distinct neural correlates and connectivity. *NeuroImage*.

Hancock, P. J. B., Bruce, V., Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, 4(9), 330-337.

Kroll, N., Knight, R., Metcalfe, J., Wolf, S., & Tulving, E. (1996). Cohesion failure as a source of memory illusions. *Journal of Memory and Language*, 35, 176-196.

Lindsay, R. C., & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, 70, 556-564.

Loftus, E.F. (1993). Made in memory: Distortions in memory after misleading communications. In G. Bower(Ed.), *The psychology of learning and motivation: Advances in research and theory Vol. 30*, pp. 187-215. San Diego: Academic Press.

McGrath, M. (2011). Eyewitness Reports, Identifications, and Testimony. In *Rape Investigation Handbook* (pp. 231-249). Elsevier Inc.

P. J. Phillips, H. Wechsler, J. Huang, P. Rauss, “The FERET database and evaluation procedure for face recognition algorithms,” Image and Vision Computing J, Vol. 16, No. 5, pp. 295-306, 1998.

P. J. Phillips, H. Moon, S.A. Rizvi, P. J. Rauss, “The FERET Evaluation Methodology for Face Recognition Algorithms,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, pp. 1090-1104, 2000.

Reinitz, M., Morrissey, J., & Demb, J. (1994). Role of attention in face encoding. *Journal of Experimental Psychology: Learning,Memory, and Cognition, 20,* 161-168.

Sangrigoli, S., Pallier, C., Argenti, A. M., Ventureyra, V. A. G., Schonen, S. (2005). Reversibility of the Other-Race Effect in Face Recognition During Childhood. *Psychological Science*, 16(6), 440-444.

Valentine, T., Ferrara, A. (1991). Typicality in categorization, recognition and identification: Evidence from face recognition. *British Journal of Psychology*, 82, 87-102.

Van Belle, G., De Graef, P., Verfaillie, K., Busigny, T., & Rossion, B. (2010, April). Whole not hole: Expert face recognition requires holistic perception. *Neuropsychologia, 48,* 2620-2629.

Verde, M. F. (2010, July). Conjunction errors in recognition: Emergent structure and metacognitive control. *Journal of Memory and Language*, *63, 476-488.*

Washington, DC: National Institute of Justice. Huff, C. R., Rattner, A., & Sagarin, E. (1996). Convicted But Innocent: Wrongful Conviction and Public Policy. Thousand Oaks, CA: Sage Publications.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wells, G. L. & Bradfield, A. L. "Good, you identified the suspect": Feedback to Eyewitnesses Distorts Their Reports of The Witnessing Experience. *Journal of Applied Psychology,* 83(3), 360-376.

Wells, G. L. & Loftus, E.F. (2003). Eyewitness memory for people and events. A. M. Goldstein (Ed.) *Handbook of Psychology. Vol 11 Forensic Psychology* (I.B. Weiner, Editor-in-Chief). New York: John Wiley & Sons, pp 149-160

Wixted, J. T. (2007). Dual process Theory and Signal-Detection Theory of Recognition Memory. *Psychological Review*, 114(1), 152-176.

Wright, W. B., Sladden, B. (2003). An own gender bias and the importance of hair in facial recognition. *Acta Psychologica*, 114, 101-114.

23 Years After Improper Photo Lineup Led to Wrongful Conviction, DNA Proves Thomas McGowan's Innocence in Dallas County Rape. (2008, April 15).

Yonelinas, A. P., Kroll, N. E. A., Dobbins, I. G., Lazzara, M., & Knight, R. T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember/know, process dissociation, and receiver operating characteristic data. *Neuropsychology, 12,* 1–17.

Ethan D. Zimmerman

edz108@psu.edu
(215) 896-5182

Local Address
747 E Beaver Ave
State College, PA 16801

Permanent Address
1771 Lynrose Circle
Hatfield, PA 19440

EDUCATION

The Pennsylvania State University, University Park, PA
B.S in Psychology, Neuroscience Option, Fall 2011
Schreyer Honors College
Honor's Thesis: R*ecognition memory for identical, morphed, and new faces*

The Pennsylvania State University Emergency Medical Services, University Park, PA, December 2010
Pennsylvania State Certified Emergency Medical Technician – Basic (EMT - B)

ACADEMIC AND PROFESSIONAL EXPERIENCE

2011- Delivery Driver, Kildare's Irish Pub, State College, PA

- Take delivery orders and delivery food to various patrons throughout State College, PA
- Must be prompt and know how to multitask, along with knowing various routes for the most efficient delivery

2011 - Plasma Center Technician (Phlebotomist), Biolife Plasma Services, State College, PA

- Greet donors as they enter and exit the building
- Answer phones within reasonable timeframe
- Take and record donor pulse, blood pressure, and temperature measures and monitor electronic donor questionnaire system
- Enter donor information into the Donor Information System (DIS)
- Assemble collection containers for plasma donation
- Stock supplies, break down empty cartons and assist with proper disposal
- Install, prime, and disconnect disposable sets on the plasmapheresis machines
- Perform venipuncture

2011 - Emergency Medical Technician, Centre Lifelink EMS (Volunteer)

- Emergency and non-emergency care and transport of the sick and injured

2008 - Research Assistant (Volunteer)
The Cognitive Aging and Neuroimaging Lab, Department of Psychology, Penn State
Research Conducted under Nancy A. Dennis, PhD.

- The lab focuses on cognitive aging in younger and older adults
- Utilize data organization and behavioral analysis using Microsoft excel and SPSS
- Some categorization and analysis of neural activity from functional magnetic resonance imaging data using MATLAB

- Administer cognitive assessment tests to older and younger participants, score such tests, recruit older and younger adults, create stimuli for studies, conduct literature searches, and help train new research assistants
- Administration of participant health screenings for fMRI experiments

2010 - 2011 Mental Health Technician, The Meadows Psychiatric Center

- Monitored inpatients on a locked unit in a psychiatric hospital
- Worked on various psychiatric units, including children's (primarily), adolescent, and adult
- Documented on patients daily activities, vital signs and changes in behaviors
- Facilitated psychotherapy groups on the individual units
- Was responsible for helping train of new employees

2007 - 2009 Dietary Aide, Peter Beck Community Retirement Center

- Worked as a waiter and host in the restaurant, a cashier in the café, and worked in utility (meal preparation, washing dishes/pans)
- Helped train new employees

HONORS AND AWARDS

2009 - 2011 Member of the National Honor Society of Collegiate Scholars
2009 - 2011 Member of Phi Kappa Phi
2009 - 2011 Schreyer Honors College Scholar
2008 - 2011 Dean's List every semester
2010 Presidential Sparks Award
2009 Presidential Freshman Award
2009 Received A+ Letter award for excellence in experimental chemistry

POSTER PRESENTATIONS

2011 Undergraduate Research Exhibition, The Pennsylvania State University, *Recognition memory for identical, morphed, and new faces*

TEACHING EXPERIENCE

2011(Spring) Teaching Assistant for Emergency Medical Technician (KINES 403), The Pennsylvania State University

SKILLS

- Microsoft Word, Excel, Access, Power point
- Adobe Photoshop, Dreamweaver
- Some experience with SPSS and MATLAB, and Fiji (ImageJ)

ATTRIBUTES

- Punctual, hardworking, determined, team-oriented, flexible, can multitask, have good communication skills, quick learner, among others

ACTIVITIES

2011 - Tutor, The Pennsylvania State University

- As of now, I tutor one student in general biology and general chemistry at Penn State

2010 - 2011 Penn State Global Medical Brigades, member

- Provided medical care to citizens of Honduras
- Raised money to purchase medical supplies and to send medical professionals along

2010 - Starfinder Mentor

- Starfinder: A program that combines athletic, educational, and personal enrichment development for young people in the Philadelphia region
- Some members of the Starfinder program (typically sophomores and juniors) visit Penn State for a weekend each year. During the year, we give them a tour of the campus, show them all the opportunities (educational and athletic) that are available at Penn State. We also provide them with information regarding majors, financial aid, scholarships, SATs, and career interests. Above all, we give them a small picture of what life as a college student could be and the advantages of being a college student.

2009 - 2010 Mount Nittany Medical Center Volunteer, Emergency Department and Surgical Center

- Transported patients to get CT scans, X-Rays, Ultrasounds, and MRIs
- Made up / prepared litters
- Transported medications, specimens, and medical records
- Provided patients and family members with meals and snacks.
- Helped train new emergency department and surgical center volunteers

2009 - 2010 Mini-Medical School Participant, Drexel University, Hahnemann Hospital

- Spent two weeks total shadowing physicians and physician assistants in both the orthopedic and neurosurgery departments for a total of approximately 100 hours
- Was able to see patients in the physician's office hours and in the operating room, including scrubbing in on selected procedures
- Was able to see selected procedures in the OR and I learned more about the medical school process, residency, and the MCAT.

2008 - 2010 Penn State Dance Marathon (THON), Springfield THON, member

- THON, largest student-organized philanthropy in the world, raising millions during the 48 hour Dance Marathon for The Four Diamonds Fund at Penn State Children's Hospital in Hershey, PA
- Participated in fundraising, social events, and spending time with the families

REFERENCES

- **Danielle Stephens,** 313 State St. Curwensville, PA 16833, (814) 592 – 9986, Former Nursing Supervisor at The Meadows Psychiatric Center

- **Beverly Pyle,** 844 Bull Run Rd. Rebersburg, PA 16872, (814) 880 – 7722, Charge Nurse at The Meadows Psychiatric Center

- **Marian Kovach,** 202 S. Centre Street, Philipsburg, PA 16866, (814) 342 – 1177, Charge Nurse at The Meadows

- **Nancy Dennis**, PhD., 513 Moore Bldg. University Park, PA 16802, (814) 865 - 1712, Research Lab Professor, Honors Thesis Advisor

- **Simon Vandekar,** 332 Vairo Blvd. Apt K State College, PA 16803, (814) 880 – 7722, Neuropsychology Lab Manager

- **Adam Piper,** 10 Vairo Blvd. State College, PA 16803, (717) 994 – 3552, Former Co-worker and current CNA at Mt. Nittany Medical Center

- **Sunie McCallen,** 10 Vairo Blvd. State College, PA 16803, (724) 989 – 1737, Former Co-worker and current RN at Mt. Nittany Medical Center