INDIVIDUAL DIFFERENCES IN INTERSENSORY INTEGRATION AND INHIBITION DO NOT PREDICT PERFORMANCE ON THE DRM PARADIGM

Thomas Jenkin

A thesis submitted for the degree of Master of Science at the University of Otago,

Dunedin, New Zealand.

Supervisor: Professor Harlene Hayne

Submitted: 16th Dec 2008

Acknowledgements

First and foremost I would like to thank my supervisor Harlene Hayne for allowing me to do this project and for her constructive guidance over the year. I would also like to thank Deryn Strange for going the extra mile in constructing materials and especially for working through the statistical analysis and the results section with me. Thank you to Juilen Gross for sound advice and Jackie Clearwater for putting up with me in her office. Thanks to Natalie Wyatt for feedback on what looks ugly in a thesis. Thanks to all my family and a giant thank you to my beautiful fiancée Brydee Strang for your support throughout this year (and all the others.)

Abstract

False memory research is a highly controversial area, particularly in the practical arenas of eyewitness testimony and memories recovered in therapy, where cases of murder and historic abuse have been overturned due to testimony being deemed inaccurate due to the influence of false memories (Brainerd & Reyna, 2005). Thus much of the academic focus on false memories has been slanted toward the practical issues such as: the authenticity of, and situational factors that cause, false memories. However, in order to fully understand the phenomenon, there is a strong need to develop robust theoretical explanations for false memory. The purpose of the present experiment was to examine whether individual differences in susceptibility to false memories could be explained, in part, by individual differences in intersensory integration, or inhibition. To address this question we compared participants' performance in a false memory task -the DRM paradigm- and their performance on two audio-visual illusion tasks -the McGurk and the Illusory Flash Effect task-, and their performance on the Stroop Inhibition task. In the DRM task, participants were presented with lists of words, which were all related to one critical lure word. The frequency with which participants subsequently recalled or recognised the non-presented critical lure was used as a measure of false memory. In the McGurk task participants were asked to report what they heard while watching a mouth speak nonsense syllables; on inconsistent trails, the spoken syllable differed from the audio syllable. In the Illusory Flash Effect task participants watched a white circle flash a computer screen; circle flashes were accompanied by audio beeps, on inconsistent trials there was on circle flash and two beeps. In the Stroop Inhibition task participants were asked to read the font colour of a list of words out loud; on inconsistent trials the font colour was different from the colour word. Overall, there were no correlations between participants' false reports of critical lures in the DRM task, and

their responses on inconsistent trials of the McGurk or Illusory Flash Effect tasks. There was also no correlation between participants' false reports of critical lures and the difference in time between responses on consistent and inconsistent trials of the Stroop Inhibition task. On the basis of these data, I conclude that individual performance in the DRM task is not predicted by individual differences in performance on measures of intersensory integration, or inhibition.

Table of Contents

List of Figures

List of Tables

Few topics in psychology are as controversial as the topic of false memory, particularly whether false memories might emerge during therapy. Brainerd and Reyna (2005) describe numerous court cases in which therapists, psychiatrists, and other mental health professionals have been sued for millions of dollars because their therapeutic practises lead their patients to develop false memories of abuse. Hundreds of other convictions, including murder, have also been overturned because the testimony provided by a key witness has been deemed the result of false memory (Brainerd & Reyna, 2005). While this controversy surrounds the practical aspects of false memory, there is a strong need to understand the cognitive mechanisms underlying how false memories develop. That is the main focus of this thesis.

False Memory

Although we all recognise that our memory is not perfect, false memories are a step beyond basic forgetting or general memory fallibility. *False* memory describes the genuine and often confident belief in experiences or aspects within an experience that simply did not occur (Brainerd & Reyna, 2005). Over the past 30 years, a range of laboratory paradigms have been developed to study the phenomenon of false memory. For example, Loftus and Miller (1978) developed one of the first procedures designed to influence participants' memory of an event. Loftus and Miller conducted five experiments (recruiting over 1,200 participants) in which participants were required to watch a slide-show of a car accident involving a pedestrian. After viewing the event, participants completed a questionnaire that contained misleading information. For example, participants were asked a question about a yield sign when a stop sign had been visible during the slide-show. Participants in the control group completed a questionnaire that contained information that was consistent with information presented in the slide-show or with only generic details included

(Experiment 2, Loftus and Miller, 1978). After watching the slides and completing the questionnaire, participants completed a memory task in which they were required to select the presented slide from two alternate slides. Across all five experiments, participants who completed the misleading questionnaire tended to select non-presented slides based on misleading information whereas participants in the control group(s) did not. Essentially, the data from Loftus and Miller (1978) indicate that participants' memories were influenced by providing misleading information after the event. This phenomenon is now commonly referred to as the misinformation effect.

Since the seminal demonstration of the misinformation effect, there have been a myriad of misinformation studies, demonstrating that post-event information can lead participants to report incorrect information about aspects of a witnessed event. This incorrect information is sometimes referred to as false memories (see, for example, Gerrie, Garry & Loftus, 2007). Gerrie et al. describe a variety of methods for introducing post-event information that have been shown to lead individuals to report incorrect information (arguably because the misinformation creates a false memory). For example, Loftus and Palmer (1974) found that manipulating the verb in the question: 'how fast were the cars going when they _____ into each other?' affected participants estimates of car speed. More specifically, participants who were exposed to the verb *smashed* tended to report broken glass, which was not present in the video. Other studies have shown that participants will report buildings in what were empty landscapes (Loftus, 1996), incorrectly describe concrete details of simulated robberies (Beli, 1989; Tverksky & Tuchin, 1989) and report incorrect colours of objects (Sutherland & Hayne, 2001), on the basis of false information encountered after the real experience.

In the original Loftus and Miller (1978) procedure, participants' memory of an event was altered by information encountered after it took place. Loftus and Pickrell (1995) conducted one of the first studies where entirely false memories were created on the basis of misinformation. Undergraduate students were provided with a written description of four childhood events, three of which had been described by an older relative, and one fabricated description which described the participant getting lost at age five. Over a 4-week period, participants were asked to describe what they could remember about each event. Loftus and Pickrell found that 25% of the participants reported a memory for the fabricated event during the verbal interviews. This finding has been replicated more recently, using photographs (Wade, Garry, Read & Lindsay, 2002). In Wade et al.'s procedure, participants were exposed to a set of photographs of themselves as children; one of which showed them participating in an activity (a hot air balloon ride) that never actually occurred. The photos were made possible by combining a genuine photo with a dummy photo, using PhotoShop. Wade et al. recruited the family members of 20 participants. Using a similar procedure to Loftus and Pickrell, each participant was interviewed three times over a 14-day period, about what they could remember about each of the target events. Wade et al. found that by the third interview, half (10) of the participants had developed false memories about the balloon ride event.

Collapsing across the literature, Strange, Hayne and Garry (2008) reported that, on average (using a weighted average), 37% of adult subjects typically develop full-blown false memories for a variety of unique childhood events (such as spilling punch over guests at a wedding or being a victim of an animal attack). Obviously, this style of research is strongly bound by ethics, however, it appears that essentially any plausible (and potentially implausible) fabrication, can be encoded as a true memory, by a substantial minority of participants.

One of the criticisms of the misinformation effect and memory implantation techniques is that it is difficult to totally remove the possibility of alternative explanations. For example, participants in Loftus and Miller (1978) may have been mislead by the questionnaires and confused, but they may not have developed *false* memories per se. Participants in Wade et al.'s (2002) study may have felt pressured to supply details about the photographs and may have fabricated memories rather than developing a true false memory. There is, however, another paradigm that does not suffer from these interpretive problems.

The Deese, Roediger and McDermott (DRM) word-recall procedure was originally published by Deese (1959), and developed further by Roediger and McDermott (1995). This procedure provides a robust and reliable paradigm for studying false memory. To complete the DRM, participants are required to listen to a list of words that are all related to a single, non-presented word. While recalling these words, participants frequently recall, or recognise, the related word that was never actually presented to them. Put another way, they develop a false memory for hearing the related, but never presented, word.

The Deese, Roediger McDermott (DRM) Paradigm

In Deese's (1959) original study, participants listened to a list of words, then immediately verbally recalled any of the words they could remember to the experimenter. While ostensibly a simple recall test, an interesting aspect of the results was that many of the presented word lists prompted *false* recall of words that were not presented. These intrusions were typically words that were semantically related to the words presented, for example words such as drowsy and bed tended to prompt the intrusion word sleep. Roediger and McDermott (1995) recognised the potential of Deese's (1959) procedure for studying false memory in a laboratory setting, and developed the DRM paradigm.

In Roediger and McDermott (Experiment 1, 1995), undergraduate students were presented with word lists similar to those used in Deese's (1959) study. During the recall test participants were required to write the words they remembered on paper. Roediger and McDermott selected word lists with the highest percentage of intrusions from Deese's original paper. After participants recalled words from each list, Roediger and McDermott also presented participants with a recognition test. The recognition test contained 12 presented words, six intrusion words, referred to as critical lures and 24 non-presented words that were unrelated to lists. To complete the recognition task, participants indicated whether they recognised each word by using a 4-point rating scale; 4 (sure the word was old/presented), 3 (word was probably old), 2 (probably new/non-presented) and 1 (sure the word was new).

Roediger and McDermott (1995) found that participants had a .65 probability of recalling presented words, and more importantly, a .40 probability of recalling a critical lure (a non-presented, but associated word). There was a .14 probability of participants recalling another non-presented (non-critical lure) word. In the recognition part of the study, Roediger and McDermott found that 75% of the presented words were given a correct rating of 4 (sure the word was old). For non-presented words, 58% of the critical lure words were incorrectly given ratings of 4, only 2% of the unrelated non-presented words were incorrectly rated 4. This trend of critical lures being falsely remembered at a rate higher than non-related intrusion words was replicated in Experiment 2 of the Roediger and McDermott (1995) study.

Roediger and McDermott's (1995) findings have been replicated in recent DRM research conducted in other laboratories (Sugrue & Hayne, 2006; Blair, Lenton & Hastie, 2002; Peters, Jelicic, Hass & Merckelbach, 2006; Peters, Jelicic, Verbeek & Merckelbach, 2007; McNally, Lasko, Clancy, Macklin, Pitman & Orr, 2004). Furthermore, even when participants are explicitly told about false memories and the purpose of the DRM memory task, they still report the critical lures (McDermott & Roediger, 1998). Because the DRM is a robust paradigm and a relatively simple procedure to conduct, the DRM has been utilised in a large variety of studies. For example, some researches have found a positive correlation between DRM performance and; executive function (Peters et al., 2006) and working memory (Peters et al., 2007); other researchers have found positive correlations between DRM performance and personality traits such as dissociation and absorption (McNally et al., 2004). In addition, Blair et al. (2002) assessed the reliability of the DRM paradigm by using a test-retest procedure over a two-week period. Participants' critical lure reports were consistent across the two-week period suggesting that the DRM measures a consistent trait of false-memory susceptibility in individuals. This list of DRM research is clearly not exhaustive, it does, however, represent a snapshot of some of the important lines of DRM research. The DRM is a vital paradigm at the heart of false memory research predominantly because it provides an objective and robust method of measuring false memory within a laboratory setting.

Current Theoretical Explanations of DRM

Given that the DRM false memory paradigm is a robust phenomenon, it is important to explore the theoretical explanations that have been developed to account for the development of false memories. At present, three well established theoretical accounts have been developed; The Activation-monitoring account, Fuzzy Trace Theory and the Constructive Memory Framework.

Activation-monitoring theory.

Roediger and McDermott (1995; 2000) have hypothesised that critical intrusion words are mentally activated due to their association with presented words and subsequently encoded into memory. Roediger and McDermott (2000) also suggest that a reality or source monitoring error, (essentially an error in remembering the source of a memory Johnston & Raye, 1981), must occur in order for a participant to report a critical lure as a presented word.

The activation component of activation-monitoring theory is based on ideas presented in Collins and Loftus (1975), who suggest that the activation of a specific word leads to activation of related semantic concepts (for example activation of the word bed will also activate the word sleep). The activation of other concepts depends on their strength of connection or relatedness to the primarily activated concept. In the DRM, a list of related words are presented to participants and the activation-monitoring theory suggests that each time the participant hears a presented word, the critical lure is cognitively activated due to its connections with that presented word. Deese's (1959) original finding that critical intrusions increased as a function of the number of related words included in presented lists, provides evidence to support the idea that spreading activation of concepts is a critical part of false memory formation.

More recently, McDermott (1997) provided support for the activation component of the explanation showing that after listening to a word list participants are primed with the critical lure word. When participants are exposed to DRM word lists they also respond to word completion tasks (that involve the critical lure) more quickly than participants who are not exposed to the word list. For example, when participants were exposed to a word list associated with sleep, they completed the fragmented word sl__p faster than participants who did not hear the list. Roediger and McDermott (2000) suggest that McDermott's findings are due to the critical lure being activated through association and priming the participants to complete the fragmented words.

The monitoring component of activation-monitoring theory is based on the concept of reality monitoring (Johnston & Raye, 1981). Johnston and Raye describe reality monitoring as a cognitive process of distinguishing whether a memory is produced from an external source (actual events) or internal/cognitive sources (such as imagination). Reality monitoring fits under a broader category of source monitoring. Logically, it stands that in order for individuals to report a false memory, some sort of reality or source monitoring error must occur as it seems unlikely that individuals would report false memories if they were aware that the source of the memory was not a real event. The empirical evidence suggests, however that manipulating reality monitoring does not necessarily predict changes in false memory formation. While source and reality monitoring errors may be common (Hicks & Marsh, 1999; Johnston, & Raye, 1981), and logically some sort of error in retrieval must occur for an individual to have a complete false memory (individuals would not claim they recalled a word, or report childhood events as occurring if they were aware that the source of the memory was some internal source, rather then an actual experience). There is evidence that false memory occurs irrespective of source monitoring manipulations (Hicks & Marsh, 1999).

In a series of five experiments, Hicks and Marsh (1999) examined whether manipulations in source monitoring could effect undergraduate participants' reporting of critical intrusions. Across the five experiments, Hicks and Marsh utilised multiple sources for the word lists including; male and female voice recordings, participant-completed anagrams, and written visual presentation of word lists. Hicks and Marsh also manipulated participants' awareness of source monitoring, and used varying source monitoring measures. Hicks and Marsh found that when completing a DRM recall task (Roediger & McDermott, 1995) with multiple sources of word lists (i.e. voice recording of word lists and self-generation of word lists) there was no significant difference in, the number of critical lures reported by participants who were aware of a simple source monitoring task (to be administered after the recall test), and those that were not aware of the task (Experiment 1). There was also no significant difference in the number of critical lures reported by participants with a source monitoring task added to the free recall aspect of the DRM (participants were required to recall words organised into source categories) and those without a source monitoring task (Experiments 2a & 2b). Hicks and Marsh did find that in some experiments (Experiments 1 & 3), presenting participants with multiple sources reduced the number of critical intrusions that were reported, however, in other experiments (Experiments 2a, 2b & 4), they did not find significant differences between the single source and multiple source groups. While this evidence may suggest that reality monitoring is not a key part of false memory formation, it must be acknowledged that reality monitoring is a difficult variable to manipulate, and further study is required to determine the importance of reality monitoring in false memory formation.

There are both strengths and limitations of using the activation-monitoring explanation as a theoretical explanation for false memory. Deese (1959) and McDermott (1997) provide evidence that the initial activation of critical lure words is due to a cognitive activation phenomenon (Collins & Loftus, 1975), essentially explaining why participants report a word such as sleep when presented with a list of semantically-related words. However, an activation model does not provide a clear explanation of why an activated word is then encoded as a true memory by the participant, or generalise well to other instances of false memory, such as described in Loftus and Miller (1978) or Wade et al. (2002). In terms of reality monitoring, it is logically sound that an individual must make some sort of reality or source monitoring error when reporting any false memory. However, empirical evidence (Hicks & Marsh, 1999) suggests that source monitoring is not an integral part of false memory, and perhaps should be considered more of an enabling factor for false memory formation rather than a vital theoretical aspect.

Fuzzy Trace theory.

Fuzzy Trace Theory (FTT), developed by Reyna and Brained (1995), is a broad psychological theory originally designed to explain how various reasoning or problem solving strategies and memory for the details of an issue interacted to form outcomes, for example solutions to problems (Reyna & Brained, 2002). Reyna and Brained (2002) describe FTT succinctly "as a model of the interface between memory and higher reasoning processes." (p. 164). The two memory concepts described in FTT are verbatim and gist memory. Verbatim memory describes memory for the surface features of a stimulus, and is usually retrieved precisely. Gist memory describes memory based on more general meanings and interpretations of a stimulus or experience. Gist memories are retrieved as general feelings of familiarity or recognition. If gist memories are encoded and/or retrieved strongly enough, it is possible that gist memories are experienced very vividly and interpreted by an individual as specific and actual memories of an event or experience (Reyna &

Brainerd, 2002). Essentially FTT explains false memories as gist memories misinterpreted as verbatim memories. For example in the DRM paradigm, participants encode the word lists verbatim but also encode the gist of the semantically related words. When participants report words, sometimes the gist memory of a critical lure is interpreted as a verbatim memory and reported as a presented word.

There is some empirical evidence to support the FTT of false memory. Seamon, Luo, Schwartz, Jones, Lee and Jones (2002) provide evidence that accurate and false recognition of words occur through different cognitive systems. Seamon et al. conducted two DRM style procedures where the word lists were presented visually on a computer screen, each word was presented individually (with a separation between each word list) in either a 20 ms or a 2000 ms exposure condition. Seamon et al. also manipulated the number of times word lists were presented to participants so that in each test some lists were presented only once while other lists were presented up to 25 times. Consistent with FTT Seamon et al. found that repeating word lists affected accurate (verbatim) and false memory (gist) in different ways; false recognition decreased with more repetitions, while accurate memory increased. This result was only found for 2000 ms words exposures however, in 20 ms exposure condition, repetition of word lists increased both correct and false recall. Reyna and Brainerd (2002) interpret these results as evidence that verbatim memory is increased by repeated exposure to word lists whereas gist memory is not. Increased verbatim memory in the repeated conditions allows participants to correct their gist memory, reducing false recognition.

The importance of Seamon et al.'s (2002) findings is that their results are perhaps not as well explained by an activation-monitoring account. Under the activation-monitoring theory, one would expect repeated exposures to DRM word lists to increase false recognition, as each exposure would further activate the critical lures cognitively. However, decreased false recognition after repeated exposures could be explained by a more accurate monitoring process. Hicks and Marsh (1999) found that increasing the number of sources (which is similar to repeated exposures) of word lists reduced false recognition. But would an increase in reality monitoring accuracy be effective enough to reduce the extra activation of critical lures due to repeated exposure? Seamon et al. argued that their demonstration that repeated exposures *reduced* false recognition and increase accurate recognition, suggests that two distinct memory systems underlie correct and false recognition. Explaining false memory formation with two distinct memory systems comprises one of the major differences between the FTT and the activation-monitoring account of false memories. Reyna and Brainerd (1995, 2002) suggest that verbatim and gist memory are two distinct memory systems, verbatim memory being fairly accurate and gist memory tending to lead to false reports, whereas the activation-monitoring account describes one memory system that encodes false information through cognitive activation.

The activation-monitoring account and the FTT account of false memories in the DRM are not mutually exclusive. The activation-monitoring account provides a plausible theory of how individuals might develop gist memories: through a cognitive activation network of semantically related concepts. Whereas the FTT explains why activated information is encoded as memory, suggesting that there is a specific gist memory system designed to encode the general meaning of a stimulus. Reyna and Brainerd's (2002) suggestion that individuals mistake gist memories for accurate and precise memories is similar to Roediger and McDermott's (2000) idea of a reality monitoring error occurring upon retrieval of a false memory.

Constructive Memory Framework.

The activation-monitoring account and the FTT both explain false memory with somewhat low-level memory functions, another possibility is that false memory is a product of higher-functions in the brain. Peters et al. (2006) propose that false memory formation is due to issues related to executive functions, describing cognitive inhibition errors as the cause of false memory phenomena.

Dodson and Schacter (2002) developed the Constructive Memory Framework (CMF) which explains false memory as arising from errors in a feature binding or integration process that occurs during memory encoding. Normally, this feature binding process functions to form a coherent representation of experiences or stimuli, however other (false) information can become incorporated, leading to false memory formation. Peters et al. suggest that feature binding errors are largely inhibition errors. Normally, the pre-frontal cortex monitors cognitive information, essentially inhibiting irrelevant cognitions or representations in favour of relevant information. An individual may report a false memory when there is a failure of the pre-frontal cortex to effectively inhibit cognitive information such as imagined events (or any other source of false memory). Johnson and Raye's (1981) reality monitoring theory is incorporated into the CMF. Dodson and Schacter suggest that during retrieval of memory the pre-frontal cortex sustains reality monitoring. As mentioned in the activation-monitoring account (Roediger & McDermott, 2000), errors in reality monitoring and required for an individual to report a false memory as true.

There is a case study (Curran, Schacter, Kenneth & Gallucio, 1997) and empirical evidence (Schacter, Buckner, Koustaal, Dale & Rosen, 1997) that the pre-frontal cortex is involved with false memory formation. Curran et al. recruited a participant who had suffered a right frontal lobe infarction, and eight control participants of matched age and years of education. The participants completed a word recognition task, where word lists were presented to participants via a computer and then participants completed a recognition test similar to Roediger and McDermott's (1995) recognition test. Participants indicated 'Remember', 'Know', or 'New' for a list comprised of both presented and non-presented words. Curran et al. found that the participant with a frontal lobe lesion had significantly more false alarms than control participants. This was interpreted as evidence for increased false memory development, as the participant who had suffered a frontal lobe lesion continued to report non-presented words as remembered and supplied autobiographical explanations of non-studied words indicated as remembered (Curran et al., Experiments 1 & 2). Control participants made significantly fewer false alarms, especially when they were required to give an explanation of their memory.

Curran et al. (1997) provide some evidence that the frontal lobe is involved in false memory formation, however their experimental participant also had damage to the thalamus and the left putamen, making strong conclusions about the link between the frontal lobe and false memory impossible. Schacter et al. (1997) completed an fMRI study that provided further evidence that activity of the pre-frontal cortex is integral to false memory formation. Participants completed a DRM recognition test, while having their brains monitored in an fMRI machine. The key comparison in Schacter et al.'s study was the difference in brain activation between correctly recognised studied words and incorrectly recognised critical lures. Both correct and false recognition was accompanied by activation in the pre-frontal cortex, however false recognition was marked by higher activation in the right anterior frontal lobe. While it is difficult to draw conclusions about cognitive process from neurological data, Curren et al. and Schacter et al. provide evidence that frontal lobe functions play a role in false memory formation.

In addition, Peters et al. (2006) provide evidence that false memory formation is related to an inability to inhibit cognitive schemas. Participants completed a DRM procedure and a Random Number Generation (RNG) task. The RNG task, originally developed by Ginsburg and Karpiuk (1994) is considered an appropriate measure of an individual's ability to inhibit previously established responses, or cognitive schema (such as generating numbers in a series). To complete the RNG task participants produce a long sequence of random numbers by adding individual values between one and ten. In order to perform well at the RNG task, participants must inhibit their natural tendencies to supply the numbers in series, repeating numbers or cycling through patterns of numbers. Williams, Moss, Bradshaw and Rinehart (2002) suggest that deficits in each area of inhibition relate to different executive functions: Repeating numbers indicates poor output inhibition, presenting numbers in a series indicates poor inhibition of cognitive schema and repeating cycles of numbers indicates poor monitoring of previous output. Peters et al. found that participants' false *recognition* in the DRM procedure was significantly correlated with their tendency to generate numbers in a series, this correlation was not found for *recall* in the DRM procedure however. Other RNG indices (repetition, cycling) were not correlated with false recognition, suggesting that false memory development may be due to poor inhibition of cognitive schemas, but not deficits in output inhibition or monitoring.

While Peters et al. (2006) provide some evidence of a link between false memory formation and cognitive inhibition, the main limitation of the CMF as an explanation of false memory is the lack of strong empirical evidence. Peters et al. only found a correlation between DRM recognition and the series measure of executive function from the RNG task. DRM recall, cycling and repetition measures of the RNG were not correlated with other measures. Potentially, the RNG is not tapping the same functions as are proposed to be involved in the CMF and more specific methodologies to measure cognitive schema inhibition are required.

One of the fortuitous aspects of false memory is that susceptibility to false memory appears to be a robust individual trait (Blair et al., 2002). One of the major lines of research in false memory is in identifying personality traits or other psychological phenomenon that are correlated with false memory susceptibility, with the hopes of elucidating the cognitive processes underlying false memory formation. Once again the DRM provides a useful paradigm for studying individual differences as critical lure reporting can be measured as a continuous variable.

Individual Differences in DRM Performance

Blair et al. (2002) provide evidence that susceptibility to false memories is a stable trait, and there is evidence suggesting that there are consistent personality variables of individuals who are more susceptible to false memories than others (Heaps & Nash, 1999; McNally et al., 2004). The most robust finding is that individuals who tend to dissociate (measured by high scores on the Dissociative Experiences Scale) and who are more susceptible to hypnotism (measured by high scores on the Tellegen Absorption Scale) show a greater tendency to develop false memories (Heaps & Nash; McNally et al. 2004).

The Dissociative Experiences Scale, or DES, was originally developed by Bernstein and Putnam (1986) who describe dissociation as: “a lack of the normal integration of thoughts, feelings, and experiences into the stream of consciousness and memory” (p. 1). More recently, Wright and Loftus (1999) found that non-clinical populations tested on the DES tended to show severe floor effects and data skewed towards zero and thus, developed the DES-C. The DES-C is presented as an appropriate scale for measuring the variance of dissociative experiences in normal non-clinical populations (Wright & Loftus, 1999). The Tellegen Absorption Scale, or TAS, measures a number of traits, including heightened sense of attended objects and imperviousness to distraction, which are related to an individual's susceptibility to hypnotism (Tellegen & Atkinson, 1974).

Heaps and Nash (1999) found that undergraduate scores in the DES-C predicted participants' tendency to report false memories. Participants completed a Life Events Inventory (Garry, Manning, Loftus & Sherman, 1996), where participants rated 42 events on how likely each happened to them before age ten. After a period of 14 or 15 days, participants completed an imagination task where they were asked to visualise some of the events included in the Life Events Inventory. Participants were then told that the experimenters had lost their previous data, and they needed to complete the same Life Events Inventory again. DES-C scores predicted whether a participant would report childhood events in the second experimental session, that were not reported in the first session, these memories were assumed to be false.

In addition, McNally et al. (2004) conducted a study comparing the characteristics of individuals who reported alien abductions to a non-alien abduction reporting control group. The underlying assumption of this experiment was that the alien abduction memories reported by the experimental group were almost certainly false. McNally et al. found that the alien abduction group scored significantly higher on the DES and TAS than the control group.

Peters et al. (2006) suggest that there are still gaps in current theories of false memory formation. Personality scales such as the DES and TAS are helpful in describing characteristics of individuals susceptible to false memory, however these scales tend to describe traits rather than provide insight into cognitive process. Peters et al. argue that researchers should focus on identifying neuropsychological concepts that are related to susceptibility to false memory. Finding a relation between false memory susceptibility and a concept that is well understood, on a cognitive level, would help to shed light on the underlying cognitive process of false memory in order to develop theoretical explanations.

Multisensory Integration

A neuropsychological concept that may be related to false memory formation is multisensory integration. Multisensory integration describes a process whereby an individual identifies stimuli in their environment through integration (and segregation) of multiple inputs from different sources (i.e., audio and visual) (Tremblay, Champoux, Bacon & Théoret 2007). The tendency to integrate information from multiple sources can lead to perceptual illusions, such as the McGurk effect (McGurk & MacDonald, 1976). The McGurk effect is a perceptual illusion caused by inconsistency between visual and audio stimuli. In a typical McGurk procedure, participants hear simple nonsense syllables such as ‘ba’ and see a mouth either speak consistent syllables (participant hears ‘ba’ mouth speaks ‘ba’) or inconsistent syllables (participant hears ‘ba’ mouth speaks ‘ga’). Participants are required to report what they *hear*. When participants are exposed to inconsistent syllables, they frequently report a fused syllable (a mix between audio and visual stimuli) or report the nonsense syllable that the mouth spoke, rather than what was actually presented in the audio.

Tremblay et al. (2007) observed that individual susceptibility to the McGurk illusion varied and hypothesised that individual susceptibility to perceptual illusions, including the McGurk effect, could be explained by a more general tendency to engage in multisensory integration. To test this idea, Tremblay et al. compared adult participants' performance in a McGurk procedure with their performance on another multisensory integration task called the Illusory Flash Effect. The Illusory Flash Effect (Shams, Kamitani & Shimojo, 2000) is a perceptual illusion which is theoretically the inverse of the McGurk Effect. In the McGurk Effect, participants *auditory* perception is affected by what they *see*; in the Illusory Flash Effect, participants *visual* perception is affected by what they hear. To complete the Illusory Flash Effect procedure, participants must correctly report how many times a small white circle flashes on a black screen. Circle flashes are accompanied by an auditory bleep. Participants are exposed to three trial types: two consistent trials where the circle either flashes once and is accompanied by one beep, or the circle flashes twice and is accompanied by two beeps, and a third inconsistent trial type where the circle flashes once but is accompanied by *two* auditory beeps. When completing the Illusory Flash Effect task, participants are typically accurate in the consistent trials, however, in inconsistent trials, participants tend to incorrectly report seeing two circle flashes (Shams et al., 2000)

Tremblay et al. (2007) found that participants' incorrect responses to the inconsistent McGurk trials and over-reporting of circle flashes in inconsistent Illusory Flash Effect Trials were significantly correlated ($r = .6$). The importance of Tremblay et al.'s finding is that while the McGurk and the Illusory Flash Effect tasks differ in sensory modality (i.e., the McGurk effect shows visual stimuli affecting audio, whereas the Illusory Flash Effect shows audio stimuli affecting visual), individual performance in each task is related. One explanation for this relation is that there is one cognitive process underlying both perceptual illusions. Tremblay et al.'s theory of a multisensory integration process could explain the underlying cognitive process unifying the two tasks. Multisensory integration suggests that individuals vary in their tendency to integrate stimulus from a variety of modalities, or perhaps, vary in levels of error-proneness during multisensory integration.

The relation between multisensory integration and false memory formation

How could multisensory integration be related to false memory susceptibility? There are aspects of Tremblay et al.'s (2007) Multisensory integration concepts that are similar to the Constructive Memory Framework (CMF, Dodson & Schacter 2002) The CMF describes *memory encoding* as involving a feature integration process, where the brain has to integrate multiple features in order to produce a coherent memory representation. According to the CMF, false memories arise from a lack of inhibition of irrelevant cognitions. The multisensory integration theory describes the brain forming a coherent *perceptual representation* by integrating multiple stimuli from differing modalities. Following the concept of multisensory integration, perceptual illusions are caused by the brain integrating inconsistent stimuli into an overall perceptual representation. The two theories are linked by a common cognitive process of integrating information from multiple sources to complete a full mental representation, a cognitive process that could be vulnerable to error if irrelevant or inconsistent information was included in the integration process. One could argue that both false memory illusions and perceptual illusions are explained by errors in inhibitive processes designed to prevent irrelevant information being encoded into memory, and to prevent inconsistent stimulus being incorporated into perceptual representations.

The Present Experiment

The overarching goal of the present experiment was to elucidate the cognitive processes of false memory development by investigating the relation between false memory, multisensory integration and inhibition. Individual performance in the DRM, the McGurk, the Illusory Flash and the Stroop Inhibition tasks were compared in an attempt to uncover any relation between individual performance in each task.

Stroop Inhibition

Peters et al. (2006) suggest that cognitive inhibition errors are a casual factor in false memory formation. If this is the case, individual performance on a measure of visual inhibition, the Stroop Inhibition test (Stroop, 1935), could be related to individual false memory susceptibility. The Stroop Inhibition test typically involves participants identifying the colour of a list of words, in two conditions: an inconsistent condition and a control condition. In a control condition, participants simply list the colour of several nonsense, or non-colour words. In some studies, such as the present experiment, participants are exposed to a consistent control condition where words are presented with the same font colour of the colour word. Typically, participants can list the colour of these control words quickly and without error. In the inconsistent conditions, participants continue to list the font colour of words, however, the words are of an inconsistent colour to the font. For example participants see the word 'Red' and are required to report the font colour black. In the inconsistent condition participants tend to read through word lists more slowly than in the control condition due to the interference of automatically reading the colour words (Stroop, 1935).

The Stroop Inhibition task is used as a general measure of an individual's ability to inhibit cognitive activity associated with processing irrelevant stimuli (Pocklington & Mayberry, 2006). Inhibition is typically measured by an individual's time difference between control and inconsistent trials, and also the mean number of errors made in each condition. A longer time to read the inconsistent trials suggests that an individual is poor at inhibiting irrelevant information. In addition, a higher error rate in inconsistent trials suggests that an individual is prone to inhibition errors. There is a commonality between Stroop Inhibition concepts and Peters et al.'s (2006) suggestion that inhibition errors in memory construction lead to false memory; it is possible that the ability to inhibit unwanted cognitive processes is a general trait, rather than a separate process for memory and visual perception. If this is the case then one might expect to see individuals who have a large time difference in Stroop tasks (suggesting poor inhibition) will also show high susceptibility to false in the DRM task.

In summary, there are two key assumptions of the present research. First, that if false memory development occurs due to integration errors in memory formation, and the same integration process underlies the multisensory integration phenomenon, then an individual who performs poorly in the McGurk and Illusory Flash Effect tasks should also show a tendency to report high numbers of critical lures in the DRM task. Also, that if inhibition plays a central role in preventing false memories from becoming integrated into genuine memory, then individuals who perform poorly in the Stroop Inhibition task should also show a tendency to report high numbers of critical lures in the DRM task.

To test these two hypotheses I recruited adult undergraduates to complete a DRM memory task, a McGurk illusion task, an Illusory Flash Effect task and the Stroop Inhibition task. Participants' tendency to misreport syllables in the McGurk inconsistent trials and tendency to over-report circle flashes in the Illusory Flash Effect task were compared to participants' tendency to report critical lures when completing the DRM task. I also compared participants' time difference between consistent and inconsistent trials of the Stroop Inhibition task to participants' tendency to report critical lures.

Methods

Participants

The participants in this study were 80 psychology students attending 100-level classes at the University of Otago. The participants consisted of 18 males and 62 females who ranged in age between 17- and 55- years (*M* = 20.29, *SD* = 5.4). Participants could apply their experimental participation towards an internal assessment option for their course.

Apparatus and Procedure

Each participant completed the DRM (recall or recognition) task, the Stroop Inhibition task, the McGurk task and the Illusory Flash Effect task. The order of presentation of the tasks was counterbalanced into eight different task orders (Appendix A). The counterbalancing was administered to control for practise or fatigue effects across the tasks and to counterbalance the McGurk and Illusory Flash Effect tasks as per Tremblay et al. (2007).

DRM Task. The DRM Task is a laboratory paradigm designed to assess false memories. Participants are presented with lists of words that are based on an associated, critical lure word. A critical lure word is not presented in the list of words, but is conceptually related to the rest of the words on the list. When participants complete a recall or recognition task, incorrect recollection of critical lure words is considered an example of false memory.

The stimuli for the DRM task were presented on Mac iBook via QuickTime video player. The stimuli for the DRM task consisted of eight word lists, each containing 15 semantically-related words. The word lists were taken from Stadler, Roediger and McDermott's (1999) Norming Study. Each list was presented orally at a rate of one word per two seconds, with a pause of one minute between each list. See Appendix B for all critical lures and word lists presented.

The DRM task had two conditions; recall and recognition, which were administered between subjects. All participants were presented with the eight words lists (Appendix B). Participants in the recall condition listened to the eight word lists. Immediately after each individual list, participants wrote, on a blank sheet of paper, as many words as they could remember. Participants completed this recall procedure for each of the eight DRM word lists.

Participants in the recognition condition listened to the eight word lists and during the 1-minute pause between auditory presentations of lists, they completed difficult multiplication problems (see Appendix C). Once all eight word lists had been presented, recognition participants completed the recognition task (Appendix D). The recognition task consisted of a list of 40 words. Of these 40 words, eight were critical lures, 24 words were from the audio lists, and eight were new words (not included in the original audio word lists) that were not critical lures. Participants were required to indicate whether they considered each word as Old (presented in the audio lists) or as New (not presented in audio lists). If participants indicated that they recognised the word from the audio lists, they also selected Remember (if the participant had a vivid recollection of the word), Know (if the participant was sure the word was presented but lacked a feeling of remembering) or Neither (if the participant was guessing). Participants also rated their confidence from 1 (not at all confident) to 5 (extremely confident) for both Old or New and Remember, Know, Neither judgments.

Stroop Task. In the Stroop Task, participants are asked to identify the font colour of a colour word that is presented on the computer screen. Participants are slower to identify the font colour, if the colour word is inconsistent to the colour of the font.

The stimuli for the Stroop Task were presented via Microsoft PowerPoint. The stimuli consisted of a list of various colour words (in various font colours) that were presented on a screen as a complete list. There was a page of consistent and a page of inconsistent words. The consistent list of words was comprised of typed words that matched the actual font colour. The inconsistent list of words was comprised of words that were a different font colour than what was typed. See Appendix E for the consistent and inconsistent word lists.

To complete the Stroop task, participants were asked to read aloud two lists of words' actual colour to the experimenter. The experimenter observed the progress of the participant to ensure accuracy and measured the time taken for participants to read out the word list for both consistent and inconsistent conditions. The consistent and inconsistent words of the Stroop task were counterbalanced between participants

McGurk Task. The McGurk task is a visual/auditory illusion in which participants' auditory perception of nonsense syllables is disrupted by the visual presentation of inconsistent syllables. In the present experiment, we used the McGurk procedure described by Tremblay et al. (2007).

The stimuli for the McGurk task were presented on a Mac iBook via a QuickTime video player. The stimuli consisted of a video image of the experimenter's mouth speaking a syllable and a simultaneous audio presentation of a syllable. Syllable presentation began and ended with the speaker's mouth in a closed, neutral position. There was a pause of four seconds between trials during which the screen was blank. The syllables were: ba, va, da, ma or tha. On consistent trials, participants were presented with matching spoken (visual) and audio syllables. For example, the experimenter's mouth formed the syllable for "ba" and the participant heard the sound "ba." On inconsistent trials, the spoken and audio syllables did not match. For example, the experimenter's mouth formed the syllable for "ba" but the participant heard "da". See Appendix F for the order of consistent and consistent trials and the syllables presented in the McGurk task.

To complete the McGurk task, participants were asked to watch the computer screen and to report what they heard to the experimenter immediately after each syllable presentation. All participants were presented with a randomised set of 15 consistent and 15 inconsistent trials. Participant responses that matched the audio presentation were considered correct for both consistent and inconsistent trials.

Illusory Flash Effect Task. The Illusory Flash Effect consists of an audio-visual illusion in which participants' perception of a visual stimulus (a white circle) is disrupted by an auditory stimulus. The stimuli for the Illusory Flash effect task were created based on the materials of Tremblay et al. (2007).

The task was presented on a Mac iBook via a QuickTime video player. The stimuli consisted of a blank screen with a white cross, presented at the fixation point in the middle of the screen. On each trial, the cross disappeared and a white circle flashed either once or twice, and then the cross reappeared. The white circle was 45 mm in diameter, subtending 2 degrees of visual angle and had a luminance of 0.02 cd/m^2. The circle was presented on screen for 67 ms, if the trial consisted of two circle flashes, an interval of 67 ms separated each individual circle flash.

An audio beep was presented simultaneously with the flashing cross. The audio beep also occurred either once or twice. The audio beep was presented for 7 ms at 3500 Hz. If a single audio beep was presented, the beep occurred 20 ms before the first circle flash. In a trial where two audio beeps were presented, the first audio beep was presented 23 ms before the first flash, and the second audio beep was presented 67 ms after the first audio beep.

On consistent trials, a single flash was presented with one audio beep, or two circle flashes were presented with two audio beeps. On inconsistent trials, one circle flash was presented with two audio beeps. In the Illusory Flash Effect task, participants were presented with 10 single beep-single flash, and 10 double beep-double flash consistent trials. Participants were also presented with 10 double beep-single flash inconsistent trials. Consistent and inconsistent trials were presented in random order throughout 30 trials. To complete the Illusory Flash Effect task, participants were asked to watch the video and report to the experimenter, immediately after each trial, how many circle flashes they had seen. See Appendix G for the order of consistent and inconsistent trials in the Illusory Flash Effect task.

Results

The primary focus of this study was to examine the relation between participants' performance in each of five tasks: the DRM (recall and recognition), the McGurk effect, the Illusory Flash Effect and the Stroop effect. Before examining the correlations between performance on these five tasks, I first examined the data for each task individually.

DRM

Recall that the key variables in the DRM procedure are participants' recall or recognition of words that were presented on the list (correct recall) and recall or recognition of words that were never presented, but that are associated with the other words on the list; these words are referred to as critical lures. First, I will present participants' overall correct and false recall, across DRM conditions (Recall, Recognition) and then examine Recall and Recognition conditions individually.

The average percentage of correct recall and false recall (of critical lures) that participants reported are shown in Figure 1. These data were subjected to a 2 (DRM condition: Recall, Recognition) by 2 (Recall Type: Correct, False) ANOVA, with repeated measures over recall type. As Figure 2 shows, there was a main effect of DRM condition F (1,78) = 51.96, $p < .01$, $\eta_p^2 = .40$, but there was no main effect of Recall Type (Correct, False) F (1,78) = 1.66, $p = .20$. The DRM condition main effect was qualified by a significant interaction between DRM Condition (Recall vs Recognition) and Recall Type (Correct vs False), F (1,78) = 26.50, $p < .01$, $\eta_p^2 = .25$. Although there was no significant difference between the two conditions for Correct recall (Recall; $M = .60$ $SD = .09$, Recognition; $M = .63$ $SD = .09$), participants in the DRM Recognition condition reported a *higher* proportion of critical intrusions (M = .81, SD = .21) than Participants in the DRM Recall condition (M = .48, SD = .25).

These findings are consistent with those reported by Roediger and McDermott's (1995); in general, participants tend to report more critical lures in the recognition, rather than the recall, condition. In fact, for this study the false recall rate in the recognition condition is higher than the correct recall rate. This finding highlights the magnitude of the false memory phenomenon in the DRM paradigm.

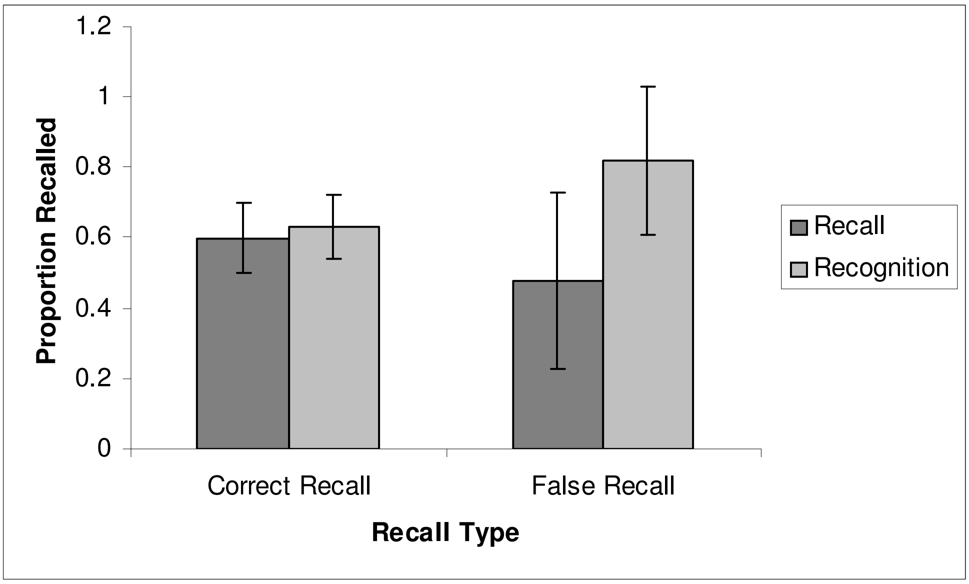

Figure 1. The mean proportion (+/- 1 SD) of correct and critical intrusion words that participants reported.

Correct Recall.

To determine the probability that participants correctly recalled the presented words, I calculated the total number of correctly recalled words for each participant as a proportion of the 120 presented words. The overall mean probability that participants correctly recalled presented words was .60 (SD = .09). Figure 2 displays the proportion of words recalled as a function of serial position of *presentation* order.

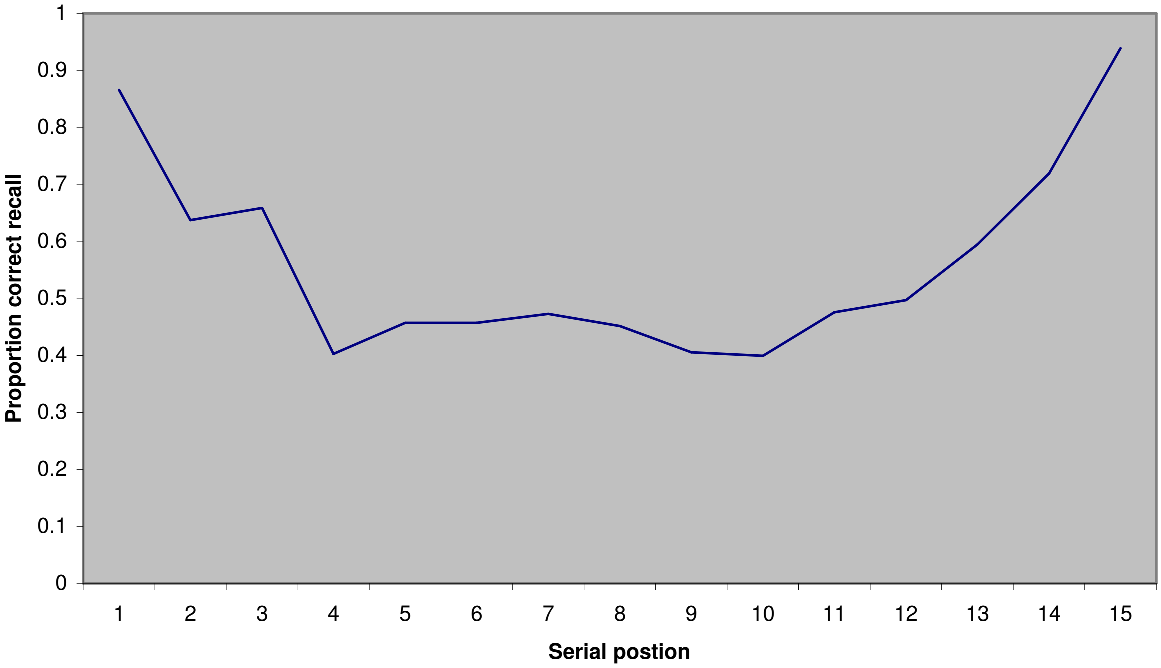

Figure 2. Proportion of correctly recalled words as a function of serial position.

As Figure 2 shows, there is a clear primacy and recency effect; participants recalled a higher proportion of words from the first and last presentation positions than from central positions. This finding is highly consistent with prior research using this paradigm (Roediger, McDermott, 1995).

False Recall.

To determine the probability with which participants recalled a critical lure, I calculated the proportion of critical lures recalled out of a possible total of eight. The mean probability that participants recalled a critical lure was .48 (*SD* = .25). To determine the probability with which participants recalled a random intrusion word (a word that was not presented *or* a critical lure word), I calculated the proportion of random intrusion words recalled out of a total of eight word lists. The mean probability that participants recalled a random intrusion word within each word list was .08 (*SD* = .24). I also calculated the mean output position of critical lures within participants' recalled lists of words, and participants' mean total word recall when a critical lure was present. The mean output position of critical lures within *participants' recalled* lists of words was 6.49 of 9.78 words (9.78 was the mean total word output of lists where a critical lure was present).

Overall, there was a significant difference in participants' recall of presented and critical lure words; participants were more likely to recall a presented word (*M* = .60 *SD* = .09) then recall a critical lure (*M* = .48, *SD* = .25), *t* (40) = 2.64, *p* = .01. There was also a significant difference in participants' recall of critical lure words and recall of random intrusion words; participants were more likely to recall a critical lure (*M* = .48, *SD* = .25) than recall a random intrusion word (*M* = .08, *SD* = .24), *t* (40) = 7.98, *p* < .01. Participants' pattern of correct recall, false recall and output position of critical lures was consistent with Roediger and McDermott's (1995) findings.

Recognition.

Recall that participants in the DRM recognition condition were required to indicate their level of memory, for each word included in the recognition test (Appendix D). Participants indicated whether they thought each word was 'Old' (presented from lists) or 'New' (not from lists), and rated their confidence in this judgement from 1 (least confident) to 5 (most confident). If participants judged a word as 'Old', they also had to indicate one of three options: 'Remember', 'Know' or 'Neither'. Participants also rated their confidence (1-5) in their Remember, Know, Neither judgment

The proportion of words judged as Old and New and the mean confidence of these judgements was recorded across Presented, Non-presented (and not Critical Lures), and Critical Lure words. The data are presented in Table 1.

Table 1. *Proportion of Old/New judgements and mean reported confidence for presented, non-presented and critical lure words (DRM Recognition). NOTE: in order to be correct participants had to judge presented words as Old, and both critical lure and non-presented words as New. Standard deviations appear in parentheses*

	Old	New	Confidence
Presented	.70 (.18)	.30 (.18)	3.89 (.38)
Non-presented	.19 (.15)	.81 (.15)	3.08 (.22)
Critical Lure	.79 (.10)	.21 (.10)	3.48 (1.37)

To determine whether participants' ratings differed across Word Type, I conducted two separate one-way ANOVAs across Word Type (Presented, Non-presented and Critical Lure). For Old/New judgements, there was a significant effect of word type, F (2,114) = 131.74, $p < .01$. Overall, Non-presented words (M = .19 SD = .15) were called Old significantly less than Presented (M = .70 SD = .18) and Critical Lure words (M = .79 SD = .10). More importantly, participants judged Critical Lures as Old as often as Presented words.

There was also a significant effect of word type on confidence ratings. F (1, 38) = 1866.06, $p < .01$, $\eta_p^2 = .53$. As a follow up, I used matched pairs t-tests to determine which Word Types had significantly different Confidence ratings (using Bonferroni correction, $\alpha = .017$). Non-presented words (M = 3.08, SD = .22) were rated significantly lower than Presented words (M = 3.89, SD = .38), t (38) = 6.51, $p < .01$, and Critical Lure words (M = 3.48, SD = 1.37), t (38) = 5.62, $p < .01$, but there was no difference in the confidence ratings that participants assigned to Presented words or Critical Lures.

Table 2 summarises the proportion of Remember/Know/Neither judgments for all words judged as Old across Word Type (Presented, Non-presented and Critical Lure).

Table 2. *Proportion of Remember/Know/Neither judgments and reported confidence for words judged as Old. Note: For non- presented and critical lure words participants have made an incorrect Old judgment. Standard Deviations appear in parentheses.*

	Remember	Know	Neither	Confidence
Presented	.65 (.20)	.26 (.16)	.09 (.12)	4.24 (.31)
Non-presented	.22 (.38)	.21 (.34)	.19 (.36)	2.78 (.64)
Critical Lure	.55 (.26)	.34 (.24)	.09 (.15)	4.12 (.23)

To determine whether word type influenced participants' Remember/Know judgements, I conducted three separate one-way ANOVAs across word type for each of the three judgement categories (Remember/Know/Neither). I found that, for Remember judgements, Presented words (M = .65, SD = .20) and Critical Lures (M = .22, SD = .26) were judged as remembered, significantly more often than Non-Presented (M = .22, SD = .38) words, F (2, 37) = 16.94, p < .01. There was no difference in the Remember judgements for Presented words and Critical Lures. For Know or Neither judgements, there was no significant effect of Word Type, (Know: F (2, 37) = 2.04, p = .14, Neither: F (2, 37) = 1.94, p = .16).

To determine the effect of Word Type on participants' Remember/Know/Neither confidence ratings I conducted a 3 factor (Presented, Non-presented, Critical Lure) ANOVA. The results indicated that participants' confidence ratings differed significantly across Word Type F (2, 37) = 26.12, p < .01, η_p^2 .59. As follow up, I used matched pairs t-tests to determine which Word Types had significantly different Confidence ratings (using Bonferroni correction, $\alpha = .017$). Non-presented words (*M* = 1.88, *SD* = .19) were rated significantly lower than Presented words (*M* = 4.22, *SD* = .19), $t\ (38) = 7.29$, $p < .01$, and Critical Lure words (*M* = 4.09, *SD* = .19), $t\ (38) = 7.32$, $p < .01$. There was no significant difference between Presented and Critical Lure words, $t\ (38) = 1.23$, $p = .23$.

Stroop

In the Stroop task, the primary variable of interest is the difference in the time it takes participants to complete consistent and inconsistent trials. Figure 3 displays the data from the present experiment. Consistent with prior research, participants took longer on the inconsistent trials relative to the consistent trials. A one way (Consistent vs Inconsistent) repeated measures ANOVA indicated that the mean difference, of 9.3 seconds (*SD* = 6.37), between consistent and inconsistent Stroop trials was significant $F\ (1, 158) = 76.11$, $p < .01$.

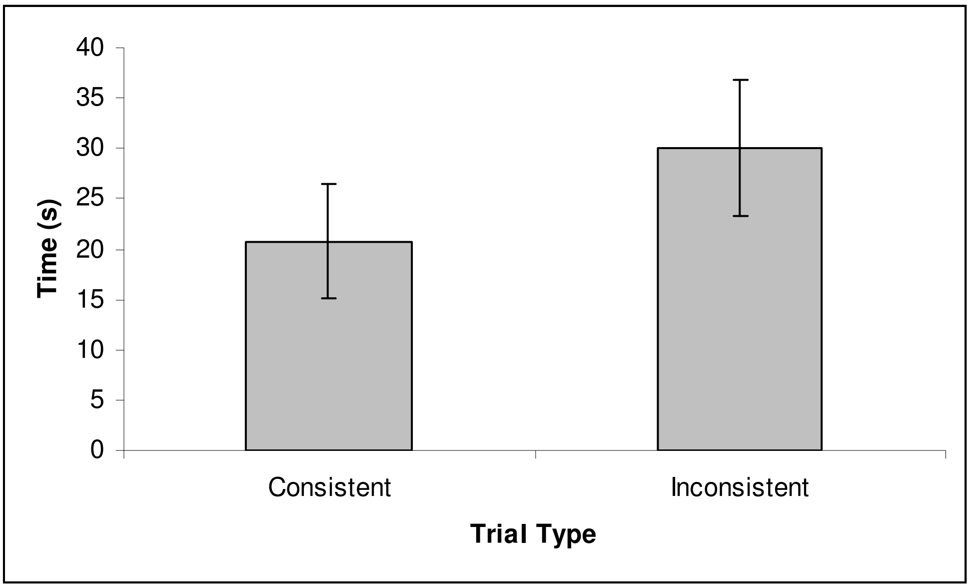

Figure 3. Time taken (s) (+/- 1 *SD*) for participants to complete consistent and inconsistent trials in Stroop Task.

McGurk

Recall that in the McGurk task, participants are required to identify presented audio stimuli while they are also viewing consistent or inconsistent visual stimuli (in the form of mouth movements). The mean percentage of participants' correct responses was calculated for consistent and inconsistent trials (see Figure 3). A one-way (Consistent, Inconsistent) repeated measures ANOVA indicated that the consistency of auditory and visual stimuli had a significant effect on participants' responses F (1, 79) = 27.86, p < .01. As shown in Figure 3, participants correctly identified the auditory stimuli more often on consistent trials (M = 93.93, SD = 10.82) than on inconsistent trials (M = 43.81, SD = 14.5).

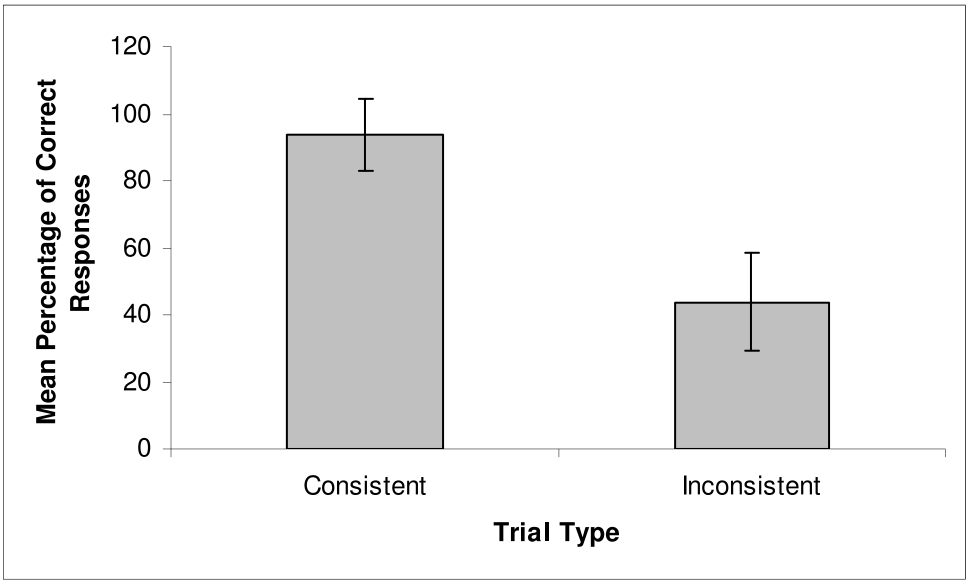

Figure 4. Mean percentage (+/- 1 *SD*) of correct responses in McGurk task for consistent and inconsisent trials.

Illusory Flash

Recall that to complete the Illusory Flash Task, participants are required to report the number of circle flashes they observe in each trial. The number of circle flashes that they see is either consistent or inconsistent with the number of beeps that they hear. I calculated the mean number of circle flashes reported across each trial type, and these data are shown in Figure 4.

On consistent trials (1,1 and 2,2) participants were very accurate in reporting the number of circle flashes they had seen. On the inconsistent trials (1,2) participants tended to report too many circle flashes (the correct response to 1,2 trials is one flash).

To determine the effect of trial type of participants' number of reported circle flashes I conducted a one-way ANOVA across Trial Type (Consistent (1,1), Inconsistent (1,2) and Consistent (2,2)). The results indicated that the Trial Type in the Illusory Flash task had a significant effect on the number of reported flashes, $F(2,237) = 450.70$, $p < .01$.

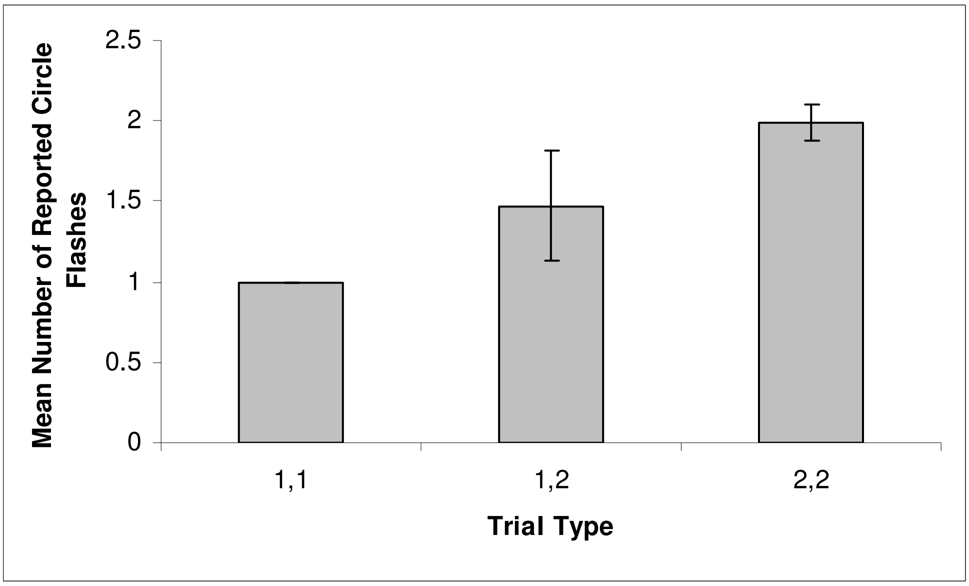

Figure 5. Mean reported circle flashes (+/- 1 *SD,* note: 1,1 trials *SD* = 0) *across* consistent single flash (1,1), inconsistent single flash (1,2) and consistent double flash (2,2) trials. Note: for 1,1 and 1,2 trials the correct number of circle flashes is one, for 2,2 trials the correct number is two.

As a follow up, I used matched pairs t-tests to determine which Trial Types were significantly different from each other (using Bonferroni correction, $\alpha = .017$). The number of reported flashes was significantly different across each trial type. In inconsistent (1,2) ($M = 1.47$, $SD = .34$) trials, participants reported more flashes than in Consistent (1, 1) trials (*M* = 1, *SD* = 0), *t* (79) = 12.81, $p < .01$. In Consistent (2, 2) trials (*M* = 1.99, *SD* = .11), participants reported more flashes than in either Consistent (1, 1), *t* (79) = 78.21, $p < .01$, or Inconsistent trials (1, 2), *t* (79) = 12.74, $p < .01$.

Correlations

Recall that the main purpose of this study was to determine whether susceptibility to false memory (as observed in the DRM) was related to susceptibility to audio-visual integration, and/or Stroop inhibition. The key variables of each task are: the percentage of critical intrusions in the DRM (recall and recognition) task; the percentage of incorrect responses to *inconsistent* trials in the McGurk task; the number of reported flashes in *inconsistent* (1,2) trials in the Illusory Flash Effect task; and the time difference across consistent and inconsistent conditions of the Stroop task. Performance in the McGurk task was measured as the percentage of incorrect responses in inconsistent trials, so that any relation between McGurk performance and other tasks would be evident as a *positive* correlation.

To determine possible relations in performance across different tasks, I performed a series of pair-wise correlations. Tables 4 and 5 display the correlations between performance in each task.

Table 3. *Correlations of participant performance in McGurk, Illusory Flash and Stroop Tasks, collapsing across DRM condition. Note: * = significance p < .05.*

	McGurk	Illusory Flash
McGurk		
Illusory Flash	-.16	
Stroop	-.24*	-.04

Table 4. *Correlations of participant performance in each Task Between DRM condition (Recall, Recognition). Note: * = significance p < .05.*

	DRM Recall			DRM Recognition		
	DRM	McGurk	Illusory Flash	DRM	McGurk	Illusory Flash
DRM						
McGurk	-.04			-.08		
Illusory Flash	.08	-.27		.20	-.26	
Stroop	.04	-.05	.03	-.04	-.43*	.05

As can be seen in Tables 4 and 5, the only significant correlation found was between performance in the Stroop task and the McGurk task. When analysed across all subjects (Table 4) a significant *negative* correlation was found, $r(78) = -.24$, $p = .05$, however when the data were analysed between DRM conditions (Recall, Recognition, Table 5), only the Recognition group showed a significant negative correlation between The McGurk and Stroop task, $r(39) = -.43$, $p < .01$.

Figures 6-10 show comparisons between the key tasks of the study. Individual participants' performances in each task are shown in each figure. The key task comparisons are: The McGurk and Illusory Flash (Figure 6), McGurk and DRM recall (Figure 7), McGurk and DRM recognition (Figure 8), Illusory Flash Effect and DRM recall (Figure 9), and Illusory Flash Effect and DRM recognition (Figure 10).

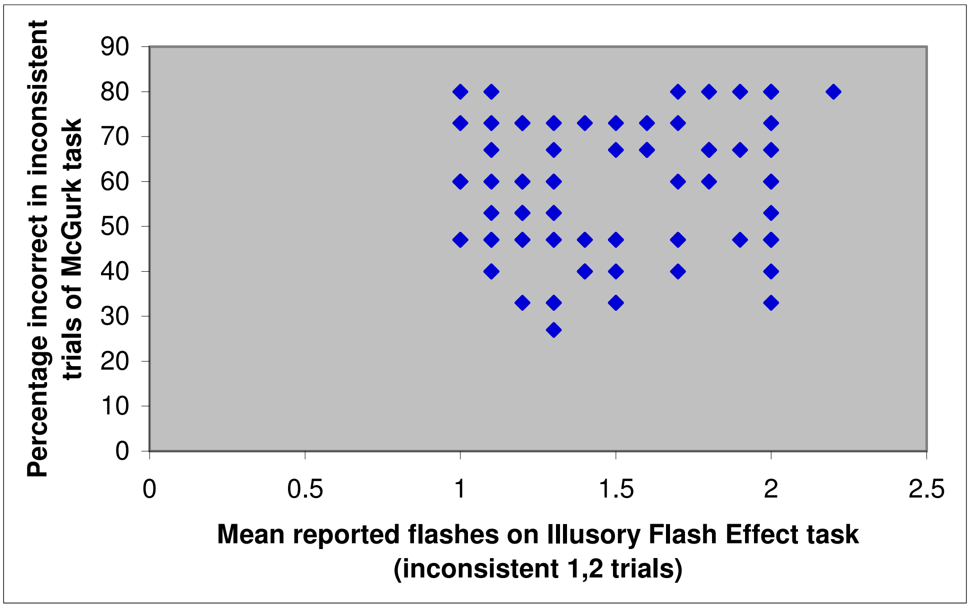

Figure 6. Scatterplot comparison of percentage *incorrect* responses in inconsistent trials of McGurk task and average number of reported circle flashes in inconsistent trials of Illusory Flash Task.

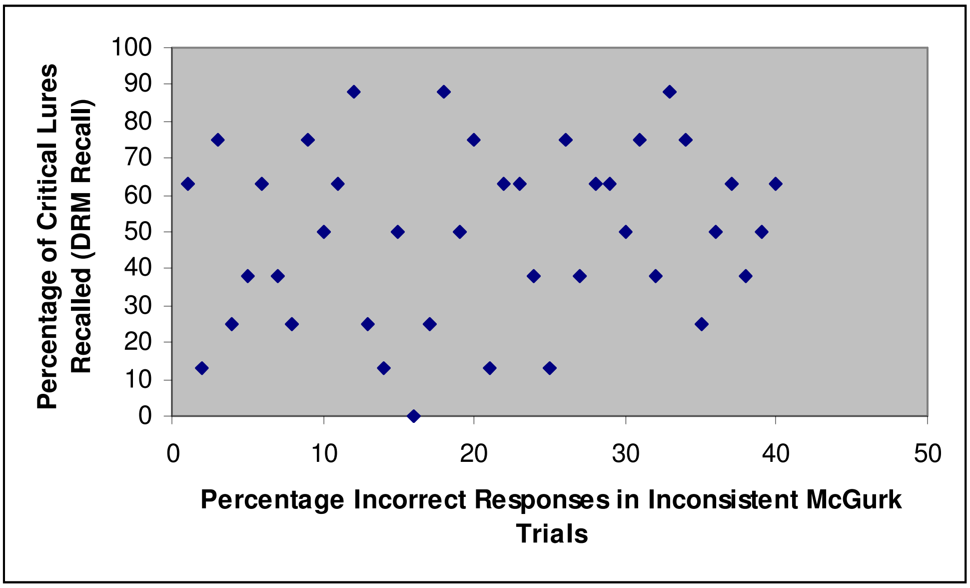

Figure 7. Scatterplot comparison of percentage *incorrect* responses in inconsistent trials of McGurk task and percentage critical lures recall in DRM recall condition.

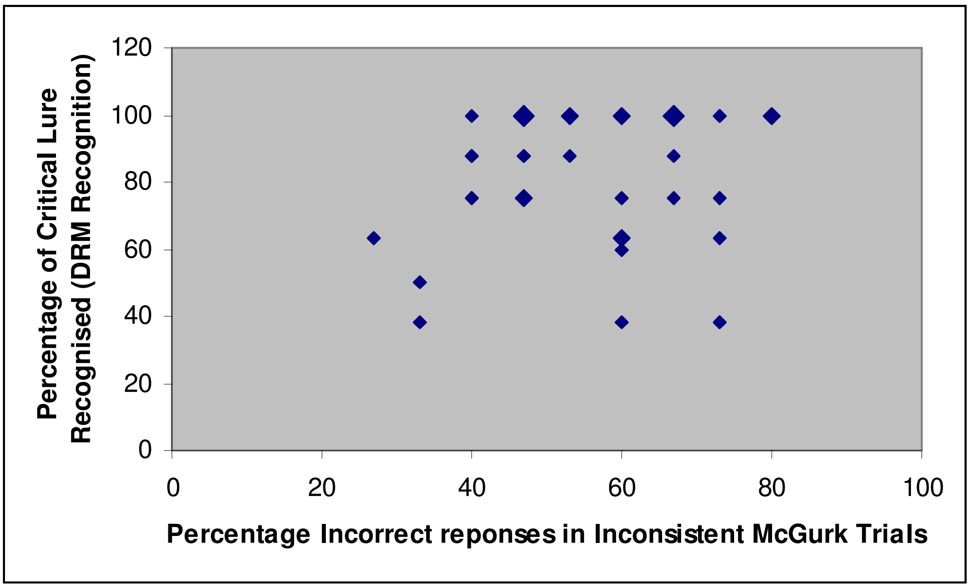

Figure 8. Scatterplot comparison of percentage *incorrect* responses in inconsistent trials of McGurk task and percentage critical lures recall in DRM recognition condition.

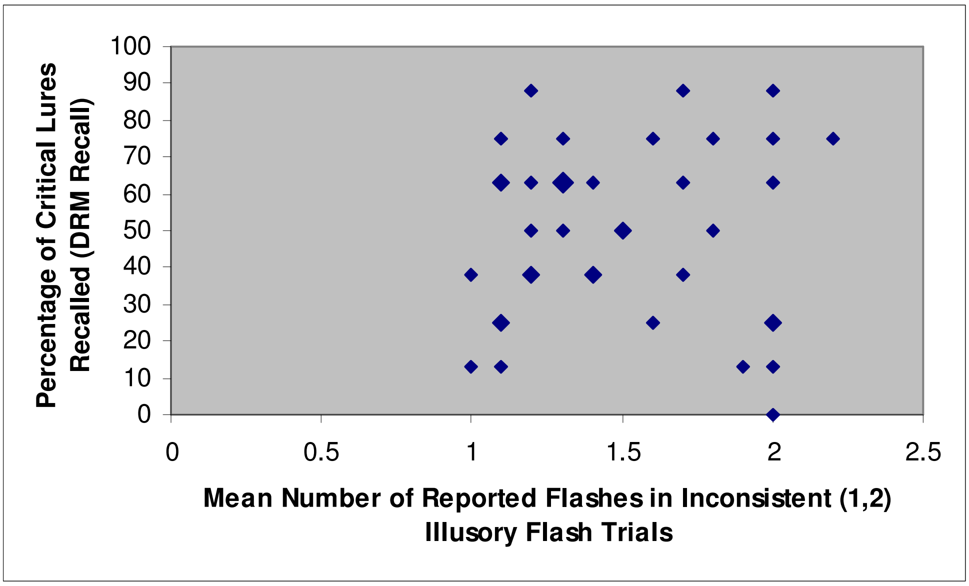

Figure 9. Scatterplot comparison of average number of reported flashes in Inconsistent (1,2) Illusory Flash trials and recalled critical lures in DRM recall condition.

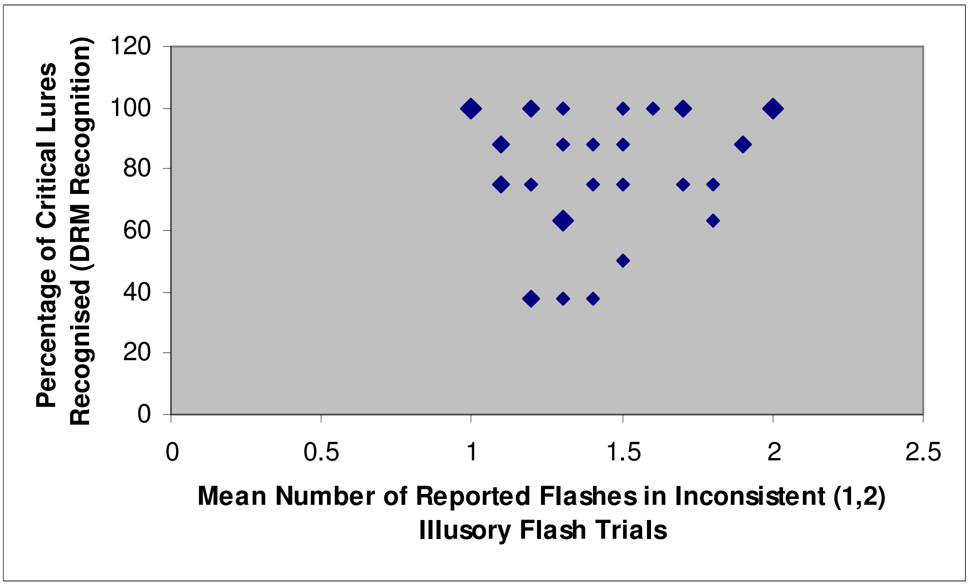

Figure 10. Scatterplot comparison of average number of reported flashes in Inconsistent (1,2) Illusory Flash trials and incorrectly recognised critical lures in DRM recognition condition.

Figures 6-10 clearly show that there is no linear relation between participants' performance on these different tasks.

In summary, the correlational analyses suggest that, contrary to our original predictions, participants' performance in the DRM tasks was unrelated to performance in the McGurk or Illusory Flash Effect tasks. In contrast to other published literature in the area (Tremblay et al., 2007), performance on the McGurk task was *not* related to performance in the Illusory Flash Effect task. Given that we did not replicate Tremblay et al.'s (2007) results, or find any other predicted correlations, the next question is why not?

Discussion

The primary purpose of the current experiment was to determine whether individual differences in the DRM task were related to individual differences in the McGurk task, the Illusory Flash Effect, and the Stroop Inhibition task. Contrary to my prediction, not only was performance on the DRM task unrelated to performance on the other tasks (see Tables 4 and 5, and Figures 6-10), I also failed to replicate the relation between performance on the Illusory Flash Effect task and the McGurk task that was previously reported by Tremblay et al. (2007). Ultimately, the only significant correlation we observed was between performance on the McGurk and the Stroop tasks. After discussing the results of each task (DRM recall and recognition, McGurk, Illusory Flash Effect and Stroop Inhibition) individually, I will discuss the inconsistency between the present findings and Tremblay et al.'s research. I will then discuss the implications and limitations of the present findings in terms of false memory research. Finally, I will discuss future areas of research for the field focussing on advancing the theoretical understanding of false memory formation.

DRM Performance

The results of the DRM recall and recognition tasks reported here are highly consistent with the original research (Roediger & McDermott, 1995) and more recent findings using similar procedures (Sugrue & Hayne, 2006 Blair et al., 2002; Peters, et al., 2006; Peters, et al., 2007; McNally, et al., 2004). As has been found before, participants frequently reported words that were not presented in a word list but were conceptually related to other words on that list. In the recall condition, participants tended to report a higher proportion of presented words than critical lures, and a higher proportion of critical lures than non-critical, non-presented words. In the recognition condition, participants misidentified a higher proportion of critical lures as Old than they correctly identified presented words as Old. Essentially, in the recognition condition participants were more likely to report false memories rather than genuine memories for words that had been presented. Also, participants were just as likely to report remembering a critical lure word (rather than just reporting familiarity or guessing) as a presented word. Participants' confidence did not significantly change when reporting a critical lure, or a presented word, as Old or selecting the Remember option, suggesting that participants reported critical lure words as confidently as if they had genuinely been presented as part of the original word lists. The fact that participants tend to report critical lures as words that they remembered vividly, and with high confidence, reinforces the notion that the false memories that are generated in the DRM paradigm are not the result of fabrications or guesswork, but are based on a genuine belief in experiences that simply did not occur.

Stroop Inhibition Task

In the present experiment, participants performed as predicted in the Stroop Inhibition Task. Participants tended to take more time to verbally report the words in inconsistent trials than the consistent trials. This is consistent with Stroop's (1935) original finding. On the basis of these findings it is argued that individuals are slower on inconsistent trials because the colour words are processed automatically, and when this processing is inconsistent with the required response, it interferes with the individual's ability to correctly identify the actual colour of the word.

McGurk Effect

The present findings on the McGurk effect are consistent with Tremblay et al's (2007) research; Participants were very accurate in reporting what they heard when the sound and the visual presentation matched (i.e., consistent trials), however, they were significantly less accurate when the two sources of information did not match (i.e., inconsistent trials). On inconsistent trials, the lip movements that participants saw apparently interfered with what they heard. Ultimately, participants reported the nonsense syllable that they saw, or a syllable that is a fusion of what they saw and what they heard.

Illusory Flash Effect

Also consistent with Tremblay et al.'s (2007) research, participants were very accurate on consistent trials of the Illusory Flash Effect task. On inconsistent trials, participants tended to report too many circle flashes, due to the interference of the auditory stimuli. Participants over-report circle flashes because the two audio beeps prompt participants to report two circle flashes, despite the fact there is only one circle flash on inconsistent trials.

Correlations

There was no evidence of a relation in an individual's tendency to report critical lures in the DRM paradigm and their performance in the McGurk and/or Illusory Flash Effect tasks. There was also no evidence of a relation between an individual's tendency or report critical lures and their time to complete the inconsistent trials of the Stroop Inhibition task. Interestingly there was also no evidence of a relation between individual performance in the McGurk and the Illusory Flash Effect task, contrary to previous research (Tremblay et al., 2007). Because the sample size was relatively large (N = 80), the lack of correlation likely reflects genuine lack of a relationship between the tasks, rather than lack of statistical power.

Contrary to our original prediction, these findings suggest that the mechanisms by which false memories are formed and the mechanisms of audio-visual illusions are created are distinct. The only significant correlation in this experiment was a negative correlation between incorrect answers on the McGurk task and time difference in the Stroop Inhibition task; furthermore, this correlation was only found for participants in the DRM recognition condition, not the recall condition. This negative relation is difficult to explain theoretically; if the relationship was positive one could suggest that both tasks were related because they both involved visual inhibition of sorts, which would be an interesting finding. However, due to the inconsistency of the correlation across the DRM recall and recognition conditions and the lack of a clear explanation, the correlation is likely an artefact of random variation rather than a reliable finding.

Multisenory Integration

As noted above, I failed to replicate Tremblay et al.'s (2007) finding that performance on the McGurk and Illusory Flash Effect tasks was correlated. In the present study no correlation was found between performance on the inconsistent trials of the McGurk or Illusory Flash Effect tasks. Although one needs to be cautious in interpreting a failure to replicate, the present results suggest that there may be no relation between the McGurk and Illusory Flash Effects. The original Tremblay et al. study was, to my knowledge, the first to compare individual performance in the two tasks, so the effect has not been replicated. Also Tremblay et al. used a relatively small number of participants (19) whereas, in this study I used 80. My results are also consistent with Tremblay et al.'s (2007) for each individual task suggesting that the illusion effects were replicated, but not however, the correlation between the two tasks.

The fact that we did not find a significant correlation could also be due to limitations of this experiment. While the methods we used for the Illusory Flash Effect were identical to those of Tremblay et al., we created new materials for the McGurk task, using more trials than Tremblay et al. and increasing the number of different syllables used from two to five. The increase in the number of permutations of nonsense syllables could have increased the overall difficulty of the task. This would mean that when an individual reported an incorrect syllable in an inconsistent trial of the McGurk task, it could have been due to the task being more difficult rather than an audio-visual integration event per se. However, our results are consistent with Tremblay et al.'s study, and one would expect poorer performance if the overall task difficulty had increased. Herein lies one of the main limitations of the McGurk task. As in Tremblay et al., researchers tend to take any incorrect answers during inconsistent McGurk trials as evidence of audio-visual integration. More care needs to be taken in assessing the McGurk task so that responses can be analysed for whether the participant reports either the syllable that they see or a fusion syllable (that is a genuine mix of the audio and visual syllables) instead of just guessing. Obviously, this sort of analysis would require strong linguistic support in order to correctly identify fused syllables. Future research using the McGurk task should only measure reports of seen or fused syllables as audio-visual integration events, rather than any incorrect answer.

In other areas, more recent research has focussed on including tactile sensation and proprioception with audio and visual tasks (Lugo, Doti, Wittich & Faubert, 2008). Lugo et al. found that EMG responses to an electric stimulus can be increased by exposure to audio or visual stimulus applied simultaneously to the electric stimulus. Essentially, participants show greater sensitivity to, or feel a stronger shock when the electric stimulus is applied simultaneously with an audio beep, or a visible spike in an oscilloscope reading.

It appears that multisensory integration as a process occurs in many different situations (i.e., the McGurk effect, the Illusory Flash effect and Lugo et al.'s, 2008) study), however, it has not established whether a generic or overarching multisensory integration process is occurring or whether each sensory modality interacts in different ways. More empirical research is needed to determine whether the general phenomenon of multisensory integration underlies the illusion present in both the McGurk and Illusory Flash Effect tasks and the effects observed by Lugo et al.

False Memory

Unfortunately, due to the lack of correlation between critical lure reporting in the DRM task, and the audio-visual illusions (McGurk and Illusory Flash Effect tasks) the present data shed little light on the underlying mechanism responsible for the development of false memories in the DRM task. One potential limitation of the present experiment was the homogeneity of the sample. The sample was entirely composed of undergraduate psychology students. While stage one undergraduate students are a somewhat diverse group, and there was a substantial level of individual variation in each task, there is a possibility that there simply was not enough variability, in each task, to accurately detect correlations between individuals' task performance. However, the pattern of results of each individual task were consistent with previous studies, therefore the results of the present study could show a genuine lack of relation (rather than revealing limitations of the experiment). Thus, on the basis of the present results we conclude that individual differences in multisensory integration do not predict individual differences in DRM performance.

Future research should focus on developing and integrating the three main theoretical explanations for the development of false memories; the Activation-monitoring account, Fuzzy Trace Theory and the Constructive Memory Framework. For example, recent research conducted by McDermott (2006), suggests that repeated testing of the DRM task results in slightly different trends for correct and false recall. McDermott manipulated the number of times participants were exposed to a DRM word list, and observed whether varying the number of initial tests effected correct and false recall in a final testing phase. In the final testing phase McDermott found that correct recall increased as a function of the number of initial tests. Recall of critical lures was higher if there were any initial tests, but there was no difference across how many initial tests were administered. As the correct and false memory measures were affected differently by the experimental manipulation, McDermott suggests that this pattern of results is consistent with Reyna and Brainnerd's (2002) theory of two distinct memory systems, gist and verbatim, underlying the formation of false memory.

Reyna and Brainerd (2002) suggest that repeated testing should increase verbatim memory recall, but not gist memory. There are still small inconsistencies however; Seamon et al. (2002) found that repeated testing of the DRM task resulted in a decrease of false recognition, whereas McDermott found that repeated testing resulted in an increase of false recall, with less increase in correct recall. Potentially, this inconsistency is due to the practical facets of recall tests versus recognition tests, such as participants being more inclined to report guesses in a recall task than a recognition task. Further replications of repeated DRM testing is needed, incorporating both recall and recognition tasks to determine the effect of repeated testing on each, and whether the difference in trends shown between Seamon et al. and McDermott are a robust effect.

In terms of the Constructive Memory Framework (CMF), Peters et al. (2006) suggest a promising line of research; focussing on linking specific neuropsychological concepts (that are already well understood) to false memory, in order to elucidate the cognitive processes that underlie false memory. Peters et al., present interesting data linking variation in false memory susceptibility to variation in the executive function of inhibiting cognitive schemas. Unfortunately, the experiment had some inconsistencies; this correlation was only found for DRM recognition conditions, not

DRM recall. Potentially the inhibition of cognitive schema is only important in preventing the false recognition of stimuli, but not every aspect of false memory. Also the present study found no relation between the individual performance in the DRM task or the Stroop Inhibition task, a common measure of executive function. There is some potential for study in this area, in particular, brain lesion studies focussed on the frontal lobe, and focusing on other executive functioning traits that have yet to be empirically tested that may correlate with DRM performance.

The activation-monitoring account provides a strong explanation for why critical lure words are reported in the DRM paradigm, suggesting that when related words are presented to an individual, the critical lure is cognitively activated also. However, under this theory, one would expect repeated testing of the DRM would elicit increased critical lure reporting. Seamon et al. (2002) found that repeated testing of the DRM did not lead to increased critical lure reporting. Also while source monitoring errors are logically needed in order for a false memory to be reported, Hicks and Marsh (1999) found that manipulations in source monitoring did not effect false memory reports, suggesting that source monitoring errors are more of an enabling factor in false memory development than a key aspect of the process.

At present it seems that, of the three major theories presented here, the Fuzzy Trace Theory (FTT) has the most support. Seamon et al. (2002) and McDermott (2006) both present evidence that accurate and false memories are affected by an experimental manipulation of repeated testing in different ways, suggesting that the mechanisms underlying accurate and false memories are different. Brainerd and Reyna (2002) suggest that these two mechanism are called verbatim and gist memory, Verbatim memory is based on exact surface features and is somewhat precise, whereas gist memory is based on the meaning and associations of a stimulus and is based on general feelings of familiarity. False memory is suggested to occur when an individual experiences a strongly encoded gist memory and interprets it as a verbatim or a precise memory.

Conclusion

Despite an ambitious theoretical argument, the results of the present study show that there are no robust correlations between individuals' performance in the DRM task, the McGurk task, the Illusory Flash Effect task or the Stroop Inhibition task. Interestingly, there was no correlation between the McGurk task and the Illusory Flash Effect task. The lack of correlation suggests that; these audio-visual illusions are not well explained by an overlying, general, multisensory integration process, but perhaps by modality specific integration processes.

Given that the issue of false memory is such a controversial topic, there is a need for sound theoretical explanations. While it is only a small advancement, the present study indicates that individual performance in the DRM task is not predicted by performance in the McGurk task, the Illusory Flash Effect task or the Stroop Inhibition task. Essentially, it seems that although our eyes and ears sometimes play tricks on us, these tricks are separate from the tricks that are played by our memories.

References

Brainerd, C. J., Reyna, V. F. (2005). *The Science of False Memory. Oxford New York: Oxford University Press.*

Beli, R. F. (1989). Influences of misleading postevent information: misinformation interference and acceptance. *Journal of Experimental Psychology: General, 118,* 72-85.

Berstein, E. M. & Putnam, F. W. (1986). Development, reliability and validity of a dissociation scale. *Journal of Nervous and Mental Disease, 174,* 727-735.

Blair, I. V. Lenton, A. P. & Hastie, R. (2002). The reliability of the DRM paradigm as a measure of individual differences in false memories. *Psychonomic Bulletin and Review, 9,* 590-596.

Collins, A. M. & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407-428.

Curran, T., Schacter, D. L., Kenneth, N. A. & Gallucio, L. (1997). False recognition after a right frontal lobe infarction: memory for general and specific information. *Neuropsychologia, 35,* 1035-1049.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58,* 17-22.

Dodsen, C. S. & Schacter, D. L. (2002). The cognitive neuropsychology of false memories: Theory and Data. In A. D. Bradley, M. D. Kopelman, & B. A. Wilson (Eds.), *Handbook of Memory Disorders,* (pp.343-362). Chichester: Wiley & Sons, Ltd.

Garry, M., Manning, C. G., Loftus, E. F. & Sherman, S. J. (1996). Imagination inflation: imaging a childhood event inflates confidence it occurred. *Psychonomic Bulletin and Review, 3,* 208-214.

Gerrie, M. P., Garry, M. & Loftus, E. F. (2007). False memories. In N. Brewer & K. D. Williams (Eds), *Psychology and Law: An Empirical Perspective* (pp. 223-253). New York: Guilford Press

Ginsburg, N. & Karpiuk, P. (1994). Random number generation: analysis of responses. *Perceptual and Motor Skills, 79,* 1059-1067.

Heaps, C. & Nash, M. (1999). Individual differences in imagination inflation. *Psychonomic Bulletin and Review, 6,* 313-318.

Hicks, J., & Marsh, R. (1999). Attempts to reduce the incidence of false recall with source monitoring. *Journal of Experimental Psychology: Learning Memory and Cognition, 25*, 1195-1209.

Johnson, M. K., & Raye, C. L. (1981). Reality Monitoring. *Psychological Review, 88,* 67-85.

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

Loftus, E. F. & Palmer, J. (1974). Reconstruction of automobile destruction: an example of the interaction between language and memory. *Journal of verbal learning and verbal behaviour, 13,* 585-589.

Loftus, E. F. & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annuals, 25,* 720-725.

Lugo, J. E., Doti, R. Wittich, W. & Faubert, J. (2008). Central processing modifies peripheral systems. *Psychological Science, 19,* 989-997.

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

McDermott, K. B. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin and Review, 4,* 582-586

McDermott, K. B. & Roediger, H. L. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language, 39,* 508-520.

McGurk, H. & MacDonald. (1976). Hearing lips and seeing voices. *Nature, 264,* 746-748.

McNally, R. J., Lasko, N. B., Clancy, S. A., Macklin, M. L., Pitman, R. K. & Orr, S. P. (2004). Psychophysiological responding during script-driven imagery in people reporting abduction by space aliens. *Psychological Science, 15* (7), 493-497.

Peters, M. J. V., Jelicic, M., Haas., N. & Merckelbach. (2006). Mild executive dysfunctions in undergraduates are related to recollecting words never presented. *International Journal of Neuroscience, 116,* 1065-1077.

Peters, M. J. V., Jelicic, M., Verbeek, H. & Merckelbach. (2007). Poor working memory predicts false memory. *European Journal of Cognitive Pscyhology, 19,* 213-232.

Pocklington, B. & Mayberry, M. (2006). Proportional slowing or disinhibition in ADHD? A Brinley Plot meta-analysis of Stroop Colour and Word Test performance. *International Journal of Disability, Development and Education, 53* (1), 67-91.

Reyna, V. F. & Brainerd, C. J. (1995). Fuzzy Trace Theory: an interim synthesis. *Learning and Individual Differences, 7* (1), 1-75.

Reyna, V. F. & Brainerd, C. J. (2002). Fuzzy Trace Theory and false memory, *Current Directions in Psychological Science, 11,* 164-169.

Roediger, H. L., & McDermott, K. B. (1995). Creating False Memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning Memory and Cognition, 21* (4), 803-814.

Roediger, H. L., & McDermott, K. B. (2000). Tricks of Memory. *Current Directions in Psychological Science, 9,* 123-127.

Schacter, D. L., Buckner, R. L., Koutstaal, W., Dale, A. M. & Rosen, B. R. (1997). Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. *NeuroImage, 6,* 259-269.

Seamon, J. G., Luo, C. R., Schwartz, M. A., Jones, K .J. Lee. D. M. & Jones, S. J. (2002). Repetition can have similar or different effects on accurate and false recognition. *Journal of Memory and Language, 46,* 323-340.

Shams, L., Kamitani, Y. & Shimojo, S. (2000). What you see is what you hear. *Nature, 408,* 788.

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory and Cognition, 27* (3), 494-500.

Strange, D., Hayne, H. & Garry, M. (2008). A photo, a suggestion, a false memory. *Applied Cognitive Psychology, 22,* 587-603.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Sugrue, K. & Hayne, H. (2006). False memories produced by children and adults in the DRM paradigm. *Applied Cognitive Psychology, 20,* 625-631.

Sutherland, R. & Hayne, H. (2001). The effect of postevent information on adult eyewitness reports. *Applied Cognitive Psychology, 15,* 249-263.

Tellegen, A. & Atkinson, G. (1974). Openness to absorbing and self-altering experiences (“absorption”), a trait related to hypnotic susceptibility. *Journal of Abnormal Psychology, 83,* 268-277.

Tremblay, C., Champoux, F., Bacon, B. A., Théoret, H. (2007). Evidence for a generic process underlying multisensory integration. *The Open Behavioural Science Journal, 1,* 1-4.

Tversky, B. & Tuchin, M. (1989). A reconciliation of the eyewitness testimony: Comments on McCloskey and Zaragoza. *Journal of Experimental Psychology: General, 118,* 86-91.

Wade, K. A., Garry, M., Read, J. D., Lindsay, D. S. (2002). A picture is worth a thousand lies: using false photographs to create false childhood memories. *Psychonomic Bulletin and Review, 9* (3), 597-603.

Williams, M. A., Moss, S. A., Bradshaw, J. L. & Rinehart, N. J. (2002). Randomnumber generation in autism. *Journal of Autism and Developmental Disorders, 32,* 43-47.

Wright, D. B. & Loftus, E. F. (1999). Measuring dissociation: comparison of alternative forms of the Dissociative Experiences Scale. *The American Journal of Psychology, 112,* 497-519.

Appendices

Appendix A: Counter balancing procedure.

Table 1A. *Eight counterbalanced task orders used in the experiment*

McGurk	**Flash**	**DRM**	**McGurk**	**Flash**	**Stroop**	**DRM**	**McGurk**
Flash	**McGurk**	**Stroop**	**Flash**	**McGurk**	**Flash**	**Stroop**	**DRM**
DRM	**Stroop**	**McGurk**	**Stroop**	**DRM**	**McGurk**	**Flash**	**Stroop**
Stroop	**DRM**	**Flash**	**DRM**	**Stroop**	**DRM**	**McGurk**	**Flash**

Appendix B: DRM word lists and critical lures

Table 1B. *Critical Lures (Bold) and presented words in DRM task.*

Window	**Sleep**	**Smell**	**Doctor**
Door	Bed	Nose	Nurse
Glass	Rest	Breathe	Sick
Pane	Awake	Sniff	Lawyer
Shade	Tired	Aroma	Medicine
Ledge	Dream	Hear	Health
Sill	Wake	See	Hospital
House	Snooze	Nostril	Dentist
Open	Blanket	Whiff	Physician
Curtain	Doze	Scent	Ill
Frame	Slumber	Reek	Patient
View	Snore	Stench	Office
Breeze	Nap	Fragrance	Stethoscope
Sash	Peace	Perfume	Surgeon
Screen	Yawn	Salts	Clinic
Shutter	Drowsy	Rose	Cure
Sweet	**Chair**	**Smoke**	**Rough**
Sour	Table	Cigarette	Smooth
Candy	Sit	Puff	Bumpy
Sugar	Legs	Blaze	Road
Bitter	Seat	Billows	Tough
Good	Couch	Pollution	Sandpaper
Taste	Desk	Ashes	Jagged
Tooth	Recliner	Cigar	Ready
Nice	Sofa	Chimney	Course
Honey	Wood	Fire	Uneven
Soda	Cushion	Tobacco	Riders
Chocolate	Swivel	Stink	Rugged
Heart	Stool	Pipe	Sand
Cake	Sitting	Lungs	Boards
Tart	Rocking	Flame	Ground
Pie	Bench	Stain	Gravel

Appendix C: Maths equations (completed in DRM recognition condition)

473	387	388	449
x204	x159	x128	x162
571	233	255	555
x112	x321	x194	172
280	482	193	223
x389	x69	x432	x141

Appendix D: DRM Recognition test.

Word				Confidence				
				Not at all confident			Extremely Confident	
Nose	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Sweet	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Snooze	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
House	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Cigarette	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Road	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Smell	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Table	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Legs	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Doctor	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Door	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Sniff	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Cup	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Pane	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Smoke	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Smooth	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Lawyer	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Dentist	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Tooth	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Nurse	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Window	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5

Word				Confidence Not at all confident				Extremely Confident
River	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Sister	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Rough	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Cigar	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Ready	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Sleep	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Recliner	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Blaze	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Chair	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Sugar	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Cat	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Nostril	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Slow	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Awake	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Lid	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Drive	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Anger	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Sour	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5
Bed	Old	New		1	2	3	4	5
	Remember	Know	Neither	1	2	3	4	5

Appendix E: Order and colour of consistent and inconsistent Stroop trial words. (Note: background in experiment was black)

Table 1C. *Stroop consistent trial word presentation.*

red	orange	green
green	blue	white
purple	green	orange
yellow	blue	pink
pink	white	purple
brown	pink	blue
red	yellow	green
blue	green	yellow
yellow	blue	purple
green	red	pink

Table 2C. *Stroop inconsistent trial word presentation*

red	orange	green
green	blue	white
purple	green	orange
yellow	blue	pink
pink	white	purple
brown	pink	blue
red	yellow	green
blue	green	yellow
yellow	blue	purple
green	red	pink

Appendix F: Order of syllables and consistent/inconsistent trials in McGurk trials.

TRIAL	VIDEO	AUDIO	CONSISTENT or INCONSISTENT
1	Ba	Ba	Consistent
2	Va	Ma	Inconsistent
3	Ma	Ma	Consistent
4	Tha	Va	Inconsistent
5	Da	Da	Consistent
6	Va	Va	Consistent
7	Ba	Da	Inconsistent
8	Da	Tha	Inconsistent
9	Ma	Ba	Inconsistent
10	Tha	Tha	Consistent
11	Ba	Ba	Consistent
12	Va	Ma	Inconsistent
13	Ma	Ma	Consistent
14	Tha	Va	Inconsistent
15	Da	Da	Consistent
16	Va	Va	Consistent
17	Ba	Da	Inconsistent
18	Da	Tha	Inconsistent
19	Ma	Ba	Inconsistent
20	Tha	Tha	Consistent
21	Ba	Ba	Consistent
22	Va	Ma	Inconsistent
23	Ma	Ma	Consistent
24	Tha	Va	Inconsistent
25	Da	Da	Consistent
26	Va	Va	Consistent
27	Ba	Da	Inconsistent
28	Da	Tha	Inconsistent
29	Ma	Ba	Inconsistent
30	Tha	Tha	Consistent

Appendix G: Order of consistent/inconsistent trials in Illusory Flash Effect.

NB: 2/2 = 2 beeps and 2 flashes

2 = 2 beeps and 1 flash

1 = 1 beep and 1 flash

Table 1B. *Trial order of Illusory Flash Effect.*

TRIAL	# of Beeps
1	1
2	2
3	2/2
4	2
5	1
6	2
7	2/2
8	1
9	1
10	2/2
11	1
12	2/2
13	2/2
14	2
15	1
16	2
17	2/2
18	1
19	1
20	2/2
21	2
22	2
23	2/2
24	1
25	2/2
26	2
27	2
28	2/2
29	2
30	1