The effect of testing and vocalisation on name face learning

Jade Broadhead

A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements for the degree of MRes.

January 2016

Abstract

Name-face learning is an important social function. The more name-face pairings there is to learn, the more difficult they are to remember. The aim of this research was to introduce mnemonic strategies designed to facilitate the memory of name-face pairs. Testing has been found to be a powerful tool in aiding recall of name-face pairings (Weinstein, McDermott & Szpunar, 2011) and is the first variable investigated in this study. The present study consists of two experiments, in both experiments participants were presented with 48 trials of name-face pairings which were split into four lists of 12. Participants were required to attempt to correctly recall the name of each face when tested. Experiment 1: Participants were 48 undergraduate students who were randomly split into one of three conditions; tested (tested after the presentation of each list), untested (tested after list four only) and restudy (restudied lists 1-3 and were then tested at list four). All participants were then required to complete a cumulative test on all of the faces form all of the lists they had seen. The results suggested that participants who were tested recalled significantly more name-face pairings than those who restudied and who were untested at the list four test and at the cumulative test. Vocalisation has been shown to improve memory of items (Gathercole & Conway, 1988), experiment 2 therefore included a further variable of vocalisation to assess whether vocalising names would improve recall of name-face pairings. Here, participants were 60 undergraduate students who were randomly assigned to one of three conditions; tested vocalise (tested after the presentation of each list and vocalised each name), tested not vocalise (tested after the presentation of each list and read each name silently) and restudy (restudied lists 1-3, tested at list four and read each name silently), all participants then completed a cumulative test on all of the name-face pairings they had seen from all four lists. Results from the second experiment suggested that there was no significant difference in the number of correctly recalled name-face pairings between vocalise and not vocalise participants at the list four test and the cumulative test.

Contents

List of Tables

List of Figures

Introduction

An everyday memory phenomenon people experience difficulty with is recalling the names of people they have met which can lead to embarrassing occupational and social mistakes. This is a particularly pertinent issue for the older population (Tse, Balota & Roediger, 2010) and those who are required to learn novel name-face pairs frequently. The phenomenon of proactive interference is well established in the learning domain whereby an increase in the rate of material to be learned results in increasing difficulty in learning of new material (Nunes & Weinstein, 2012).

This literature review aims to investigate why learning many names in a relatively short period of time is difficult. It also aims to investigate variables which can be used to increase the recall of stimuli. Testing, where individuals are tested at intervals throughout the presentation of stimuli will be discussed, as well as the effect of vocalisation on memory, where the effectiveness of reading aloud information will be discussed. Previous research, although investigated independently has suggested that both testing and vocalisation can improve memory. Research using word lists will be discussed and the use of name-face pairings as stimuli will be outlined as a novel area of applied research. Remembering a name associated with a face is an important every day, socio cognitive task (Parr, 2011) and therefore it is important to investigate memory in regard to name-face learning.

A person's face is the key to their identity and it is therefore important to understand how faces are remembered as well as how other information about the person is accessed through face recognition. Bruce and Young (1986) proposed a theory of face recognition which could help to explain why remembering an individual's name is difficult, especially in comparison to other person identify information. They suggested that individuals can extract several pieces of information from a person's face. Their theory outlines eight segments of facial recognition which are accessed separately. For someone recognising a familiar face these include structural encoding of the face, face recognition nodes, person identity nodes and name generation. The theory outlined by Bruce and Young (1986) states that the first component to be accessed is the face recognition unit, hence faces are recognised first before additional information about the person is remembered. This is then followed by the person identity nodes where semantic information about a person, for example, their occupation can be accessed. After this, the name generation node is accessed. This therefore suggests that an individual will be more likely to recall an individual's occupation as oppose to their name because the person identity nodes relating to such information are accessed first.

Empirical research can be used to support Bruce and Young's (1986) theory. Young et al (1993) found that when participants were asked whether a face was familiar to them their reaction times were quicker than when they were asked if the face was a politician. This supports the theory that face recognition occurs before person identity information is retrieved. In addition, Kampf, Nachson and Babkoff (2002) identified that participants reaction times for categorising familiar faces with their occupation was faster than for stating the names of the same faces. This therefore supports the notion that recognition of faces precedes the identification of individuals' occupations which precedes identifying a name for an individual. McWeeny, Young, Hay and Ellis's (1987) study involved participants learning both the name and occupation of an unfamiliar face. The same word (e.g. cook, baker, butler) was either presented as a name or as an occupation. They found that performance for the recall of names was significantly worse than recall for occupations (i.e. when ‘baker’ was used as a name it was more difficult to retrieve than when ‘baker’ was presented as an occupation).

Cohen (1990) explained this effect by proposing that it is the meaningfulness of information which determines memorability of it. A key difference between names and other kinds of person identity information is that names are relatively meaningless, arbitrary and difficult to image. The particular deficit for recall of names compared to other person identity information has been explained as names being less meaningful than other types of information about a person and so this type of information lacks semantic associations (Cohen, 1990; Cohen &Burke, 1993). Therefore, learning new name-face pairs is a paired associate task (Weinstein, McDermott & Szpunar, 2011) where items to be associated are visual in nature (a face) but difficult to verbalise (Brown & Lloyd-Jones, 2003) and items which are verbal (a name) in nature but present meaningless information (Carpenter, & DeLosh, 2005).

To assess his claims Cohen (1990) created non-words which were used to manipulate the meaningfulness of possessions and occupations to assess whether names are better remembered when further information about the person is meaningless. Non words used as possessions were completely made up (e.g., wesp). Their findings revealed that even when participants were completely unfamiliar with the made up word (in contrast to a name they would have heard before), there was no difference in participants performance to recall a name or the made up word. This is in contrast however, to participants who were exposed to real words relating to possessions (e.g., boat) where participants were able to recall the possession more so than the name. When occupations are made up (e.g. ryman) they are deemed meaningless and names are remembered more so than unfamiliar occupations. Cohen's (1990) study showed that when participants have the ability to meaningfully encode a name (when other information about the person is meaningless) they are better able to recall the name. Recall of the meaningless item, either the name or made up occupation is more dependent on the recall of the meaningful item. Cohen (1990) stated that the person identity node where meaningful information is stored is retrieved first, information which is deemed meaningless is retrieved from the person identity node only. Thus names are generally more difficult to retrieve than occupations because they are deemed meaningless. It is challenging to access names because names are not connected to the semantic network. It could therefore be argued that it is as difficult to remember the correct name of a face as it is to attach and remember a non-word to a face as both lack semantic associations.

There has been little research which contradicts Bruce and Young's (1986) theory. Brédart, Brennen, Delchambre, McNeill, & Burton,'s (2005) study, however found that response times to name faces were faster than information about their education status when participants were presented with the faces of their colleagues and were asked to either state the highest qualification they had or their name. This somewhat contradicts Bruce and Young's (1986) theory of face recognition as participants were able to name faces before they could recall specific information about their identity, however it could be argued that this was because the participants know the names of the faces very well as they were close colleagues. Similarly, Calderwood and Burton (2006) found that both children and adults recalled names faster than semantic information of very familiar faces. They explained that this could be because it is much more likely for names to be retrieved in everyday interactions with people than semantic information such as their occupations or nationalities; it is therefore advantageous to recall names rather than semantic information about the individual. It could therefore be argued that in some cases where individuals are very familiar with faces that names are generated before the person identity nodes are accessed. Additionally, Bruce and Young's (1986) theory of recognition does not explain covert recognition, where someone can recognise a person's face without awareness, which has been outlined in cases where patients have congenital prosopagnosia (Rivolta, Palermo & Schmalzl, 2013).

It is important to understand how information such as that of learned name-face pairs can be forgotten. Interference can lead to the forgetting of information, when information is retrieved memories compete with each other and some information may be difficult to access. Proactive interference can occur when there is a large amount of information which is being remembered; the learning of new information is disrupted and may therefore be more difficult to remember than information which was learned earlier (Tomlinson, Huber, Rieth & Davelaar, 2009). Proactive interference can be problematic in many areas where an individual is required to remember and recall information.

Proactive interference can have an effect on eyewitness testimonies where previously learned information has altered the view of a witness when it comes to them recalling the event (for example, the witness may know something personal about an offender and this may distort their view of what they have seen), they may therefore give a false account of what has happened (Lindsay, Allen, Chan & Dahl, 2004). In addition, the negative effect of proactive interference can be used to explain why learning a lot of information can be difficult. For example, if an individual is required to remember and recall several lists of words the learning of earlier word lists can disrupt their ability to remember later word lists (Lustig, May & Hasher, 2001 & Nunes and Weinstein, 2012). It could therefore be expected that the more information someone is required to learn, the more the performance of later learned information is hindered.

Bennet (1975) argued that the process of forgetting was due to the natural decay of memory traces, he also stated that the forgetting of information could be due to other competing memories. In addition Baddeley, Gathercole and Papango (1998) stated that forgetting arises when different memories disrupt each other. A considerable amount of research conducted in the area of proactive interference and memory has investigated the ability to retain and later recall word lists. Many researchers such as Nunes and Weinstein (2012) and Wissman, Rawson and Pyc (2011) have established that when participants learned multiple lists of words their performance was better for earlier lists than later lists. This suggests that proactive interference affected performance of participants; it could be argued that the learning of earlier word lists disrupted the participants' ability to remember later list words.

The testing effect has been shown to improve recall of items. Studies such as Pastötter, Schicker, Niedernhuber and Bäuml (2011) have conducted research where participants studied lists of words and were either tested or untested throughout the presentation of lists 1-4 and then carried out a free recall task on the list five test. Pastötter et al (2011) studied the extent to which retrieval through testing facilitates the encoding of information across each list. They found that where participants had no opportunity for retrieval (were not tested after each list) there was an increase in alpha power, and through encoding of lists one to five, however no increase in alpha power was found when participants had the opportunity for retrieval and were tested between the presentation of each word list. The results found suggested that participants who retrieved information (were tested between each list) reset their encoding processes for each list and this means that the encoding of later lists was as effective as the encoding of already learned lists which were presented earlier. Participants were therefore able to recall words they had learned from list five equally as well as words they had learned from the first list. It could therefore be expected that participants who have the opportunity for retrieval perform better as proactive interference is reduced.

Other researchers have proposed theories to explain how the testing effect promotes retrieval of information. Jang and Huber (2008) stated that the process of information retrieval between studied lists of stimuli promotes context changes which encourage the segregation of each list. They proposed that when information from each list is encoded, the memory for each item binds to the existing image of the individual's internal context. The retrieval process which occurs when an individual is tested after each list changes the internal context; this means that for each presented list there are specific context cues. Thus, when they are tested the specific context cues are enforced; this enhances the discrimination of each list and in turn reduces the effect of interference between each list (Pastötter et al, 2011). Divis and Benjamin (2014) suggested that the similarity between the learning of information and the test encourages greater retention of information. Divis and Benjamin (2014) stated that retrieval through testing encourages internal context change; therefore there is superior contextual segregation among learned items which were presented either prior or following the retrieval event. List segregation improves memory because it decreases the interference between competing events such as the information from a previously learned list competing with information in the current list test.

Many studies have suggested that participants who are tested between the presentations of information perform better than participants who are not tested between lists of stimuli. However, Roediger and Karpicke (2006) identified a major confound in the literature and argued that it is possible for participants to perform better when they are tested merely because they restudy the information while it is being presented during the testing phase and not from specifically retrieving previously learned information.

Research in the education literature has suggested that a further condition should be introduced to account for this potential confounding variable. Szpunar, McDermott and Roediger (2008) therefore included three conditions in an experiment investigating the effect of testing on memory recall of word list items. One group of participants were tested after each list of words (5 lists of 18 words), another was tested on the last list only and a third group restudied the information from lists 1-4 and were then tested on the material in list 5. They assumed that if testing was due to the restudy of information then participants in the restudy group would recall list 5 words as well as or better than those in the tested condition. However, if the act of retrieval (through testing) in the initial tests was responsible for higher performance on list 5 then it is likely that those in the tested condition would perform better than those in the untested and restudy conditions. Their results suggested that participants in the tested condition recalled significantly more correct list 5 words than participants in both the not tested and restudied conditions which demonstrated that it is retrieval processes when testing which are more important than the restudy of material when aiming to remember information.

Carrier and Pashlar (1992), Cull (2000), Schmidmaier, Ebersach ,Schiller, Hege, Holzer and Fischer (2011) and Carpenter (2009) also found similar results to the research conducted by Szpunar et al (2008). These studies found that participants who were tested between the presentations of stimuli performed better than participants who restudied the stimuli; it could therefore be concluded that testing therefore benefits retention more so than the restudying of material. Anderson and Bower (1972) stated that the reason for the testing effect contributing to better retention of information in comparison to restudying information could be due to the notion that the restudy condition involves information merely being presented to participants. The process of testing however requires an individual to activate their memory in order to retrieve a specific target piece of information.

Many studies conducted on testing and recall such as Szpunar et al (2008) and Weinstein et al (2011) have found that participants who are tested between lists of information perform better on a final cumulative test than participants who are untested. Arnold and McDermott (2013) used the test potentiated learning hypothesis to explain the results of these studies on recall, testing and restudy opportunities. They stated that the retrieval of information (through testing) benefits the subsequent learning and recall of information because participants are able to identify where they have performed unsuccessfully during their attempts at retrieval and they are then more encouraged to learn during future opportunities to study further information. Participants who restudy information, however, are not tested and will therefore not recognise how well they have performed and this may not affect the learning of subsequent information. Each time an individual makes an attempt at retrieving information, they can determine the extent to which each piece of information is remembered. If they cannot remember the information they could make better use of encoding strategies which can be used on the items which are not remembered to potentially increase their performance during opportunities for recall (Roediger & Karpicke, 2006).

Zaromb and Roediger (2010) conducted research on the testing effect using words as stimuli; they determined that testing improves the long term retention of words, in contrast to the restudying of words on a final free recall test. Zaromb and Roediger (2010) suggested that both categorical knowledge and the success of previous attempts at recall allow individuals to develop plans for retrieval at future opportunities for recall. It could therefore be argued that these two complementary retrieval schemas which derive from testing may account for the testing effect which is demonstrated through delayed free recall in this research.

Takashima, Segers, Fernándezb, Verhoevena, and van den (2013) conducted a functional magnetic resonance imaging (fMRI) study to define the difference between areas of brain activation for participants who restudied information and participants who were tested on already learned information. Tested participants performed better on a later memory recall task than participants who restudied the stimuli. The researchers found that better performance on a later test was determined by greater activation in the left middle temporal gyrus and in the inferior parietal lobe which was demonstrated by participants in the tested condition but not by participants in the restudy condition. Lau, Phillips and Poeppel (2008) stated that the inferior parietal lobe is responsible for incorporating semantic material into a greater context, and it also accesses the appropriate information in memory to select the correct piece of information (Blumenfeld & Ranganath, 2007). It could be argued that greater activation throughout the testing phase in comparison to the restudy phase contributes to the idea that testing requires more intentional and effortful processing than restudying (Takashima et al, 2013).

Takashima et al (2013) stated that regions of the midbrain, which is where areas of the brain's motivation and reward system are, are also activated during the testing phase. Participants who were tested were therefore able to highlight motivationally significant information and direct their attention toward the relevant information during the encoding of information (Shohamy & Adcock, 2010). Takashima et al (2013) argued that the memory trace could therefore have been strengthened by the increased activity in certain parts of the brain and this strengthening of associations between retrieval cues and relevant responses is responsible for tested participants performing better than participants who restudied information.

The study conducted by Szpunar et al (2008) as well as other research such as Pastötter, Weber and Bäuml (2013) and Nunes and Weinstein (2012) found that testing improved performance of recall and these studies were conducted using word lists as stimuli. It is important to investigate the testing effect in research which uses stimuli other than word lists. Agarwal, Karpicke, Kang, Roediger and McDermott (2008) studied the testing effect on both closed and open book tests. Participants studied prose passages and either restudied the material or had a closed or open book test, they concluded that taking either kind of test improved long term retention of information in comparison to restudying the material. However, in this study participants in the restudy condition predicted that they would perform better on a later recall task than participants who were in either test condition. Tested participants therefore failed to foresee the effectiveness of testing when compared to restudying information; it could therefore be argued that individuals have little metacognitive awareness of the testing effect (Karpicke, Butler and Roediger, 2009).

Although there has been extensive research examining the testing effect in word list learning domain, there is limited research investigating the utility of this effect in name-face paired learning. Research such as Weinstein et al (2011) and Helder and Shaunessy (2008) have investigated name-face recall and have suggested that testing between lists of stimuli improves recall of stimuli on a final list and a cumulative test. Pariante (1990) stated that name-face learning is a more complex form of list learning as it is a paired association task which incorporates two pieces of information which an individual is required to remember simultaneously (the image of the face and the written name). Learning name-face pairings is an everyday occurrence and research in this area could have practical applications and help those who are required to know a large number of individuals' names when in social or work scenarios.

Studies have been conducted in the area of testing and recall using names and faces as stimuli. Landauer and Bjork (1978) showed participants images of faces which were assigned a first and last name at the presentation stage. Participants were either in the restudy condition where the face was shown with both the first and last name for a second time, or in the tested condition where participants were shown the face with either the first or last name and they had to recall the name which was not presented. The authors in this study found that participants who were tested on the names performed better in a later recall test than those who restudied the names. However, it is not clear whether the recall of tested items is due to the name-face association or the name name association (participants could have remembered which last name went with the first name or vice versa instead of remembering which name went with a face). It is important to investigate whether the testing effect exits when the face is the only cue and participants are required to name the face.

Morris, Fritz and Buck (2004) investigated the performance of participants when they were either tested on or studied the learning of classmates' names. In this study, the first student said their name aloud, the second student then stated the first person's name and then added their name, and the third student said the first two students' names and then added their name and so on. Students in the tested condition were required to recall the pervious names of students and students in the restudy condition had to read a list of the previous students' names. Morris et al (2004) found that tested participants retrieved more names than restudy participants. This study provided evidence of the testing effect in an ecologically valid field experiment thus demonstrating the testing technique can be used in real world settings to increase the number of name-face pairings remembered. The tested condition in this study, required participants to recall names as well as listen to others retrieve names of previous students. Participants in the tested condition therefore received additional opportunities to study the names (Carpenter & DeLosh, 2005). Future research should therefore ensure that testing does not involve additional study opportunities for participants. Furthermore, the different modalities of the presented names were not addressed; some participants only heard other names while others heard them and read a list of them; hence the different modalities of the to be remembered information could have affected results (Gathercole & Conway, 1988).

Tse, Balota and Roediger (2010) studied the testing effect using name-face pairings as stimuli on middle aged and older adults. On a delayed recall test middle aged participants who were in a repeated testing condition performed better than middle aged participants who were in the repeated study condition. Older participants only benefited from repeated testing when they received feedback (i.e. participants stated a name and they were informed if it was incorrect). However, when the older participants were not given feedback they performed better in the repeated study condition than in the repeated testing condition. Tse, Balota and Roediger (2010) stated that the failure to find a testing effect in this case was that older adults made more errors during acquisition, when they were initially shown the name-face pairings and testing therefore meant that participants learned the incorrect pairings which they had created during the initial presentation stage. They suggested that although feedback did not reduce the number of errors at the presentation stage they were protected against errors that may have influenced the final cued recall test. Tse, Balota and Roediger's (2010) study included the same procedure to that of Karpicke and Roediger's (2006), however they argued that as their study included feedback to participants in both the tested and restudy conditions it was more ecologically valid as when individuals test their memory in everyday situations they are likely to assess whether their responses are accurate. Furthermore, memory declines with age (Tong et al, 2013) so it could be argued that the results of Tse, Balota and Reodiger's (2010) study cannot be directly compared to findings where younger people were used as participants.

Other factors as well as testing between lists of stimuli have also been considered in relation to increasing the retention of information on memory tasks. Research conducted on everyday memory has suggested that words which are vocalised are remembered more than words which are presented visually for an individual to read silently (Tell, 1971). Gathercole and Conway (1988) studied how word lists presented in different modalities (e.g. visually, auditory, verbally) affected participants retrieval. They concluded that of all the modalities, only when participants vocalised the words did they have consistent retention of the stimuli.

The effect that vocalisation has on retrieval of information has been studied in real life classroom scenarios. Rosenthal and Ehri (2011) investigated the effectiveness of learning unfamiliar words when reading passages of a book silently compared to reading aloud. They found that the learning of vocabulary was superior for participants who used the vocalised strategy compared to those who read the passages silently. They concluded that reading new words aloud, compared to reading them silently strengthened the connections between spellings, pronunciations and meanings in memory and thus, the word and what it meant were easier to recall.

Hale, Skinner, Williams, Hawkins and Neddenriep (2007) also investigated the effect of spoken aloud material on reading comprehension of school children. After reading a passage of writing (either silently or aloud) participants were given multiple choice comprehension questions, children who read the passages aloud performed significantly better than those who read them silently. These studies contrast with earlier findings which suggested that reading aloud could potentially hinder the comprehension of information because cognitive resources are primarily given to attaining phonological recordings instead of understanding the presented information (Jones & Lockhart (1919), Juel & Holmes (1981)).

It has been stated that the translation of to be remembered information from one mode to another is an explanation of how vocalisation of words improves memory (De Haan, Appels, Aleman, & Postma, 2000). Studies have indicated that the translation of stimuli from one mode to another (e.g. seeing a word and then say it aloud) during the encoding stage increases the retention of to be remembered items (Rackie, Brandt & Eysenck, 2014). Rackie et al (2014) found that participants who were required to translate information from one mode to another (who vocalised words) performed better on a recognition task than those who read the words silently and who therefore did not change from one mode of translation to another. Forrin et al (2012) suggested that vocalisation of information improves recall because individuals are aided by their own production of speech. Forrin et al (2012) concluded that speaking information aloud produced higher recognition for items regardless of the mode in which the stimuli was initially learned.

A further explanation of how vocalising increases the retention of information is the production effect outlined by Ozubko, Gopie and MacLeod (2011). Similarly to other research on the vocalisation of items, Ozublo et al (2011) found that participants who vocalised words performed better on a recall test than participants who read words silently. The participants first read a list of items silently and then read the remaining words out loud; significantly more vocalised words were recalled than the words which were silently read. The number of recalled words was compared between conditions where words were either all read silently or all read aloud. This study concluded that participants who read all words aloud performed poorer on a recall test than participants who read half of the words aloud and half of the words silently. It could therefore be argued that the greater performance of recall for vocalisation is due to the distinctiveness it provides against silently read items and not due to the task of vocalising (Markman, 2010).

The production effect therefore means that the individual has a memory link to the production of the word when it is made more distinct to other words because it is vocalised. The words which were vocalised were translated from being read silently into being spoken, the individual therefore remembers that they have produced and heard each word. This is what makes the vocalised words distinctive and means that they are better remembered compared to silently read words. It could therefore be argued that it is of greater benefit for an individual to read aloud to be remembered information only and read silently unimportant information as this would make the to be remembered information distinctive. Ozubko et al (2011) stated that the production effect had two important factors, firstly that there was greater recall of items which were distinctive of other items at the study period and that there was greater attention assigned to material which was vocalised at the study period.

Forrin, MacLeod and Ozubko (2012) argued that the mechanism underlying the production effect is the execution of a distinct, item-specific response. The magnitude of the production effect is predictable from the number of distinct unique encoding processes involved; as they increase so too does memory retention. This explains why vocalised words are remembered superiorly to words which are read silently. Silent reading entails only one encoding process whereas vocalising words involves two further processes: articulation (the execution of a motor action) and (2) audition (hearing oneself saying the word). As vocalising involves two additional distinct processes, it results in the largest memory advantage in comparison to other methods of production. This account is supported by hierarchical performance on memory tasks; the more encoding processes involved the better the memory performance with silent reading scoring lower than mouthing which in turn scored lower than reading aloud (Forrin et al, 2012).

Hourihan and Macleod (2008) also identified the production effect in their research. They found that the number of recalled words which were produced (read aloud) were not affected by instructions to either remember or forget whereas there was better recall for words which were read silently when they were instructed to remember as oppose to forget. This supports earlier research conducted by Golding, Long and MacLeod (1994) which suggested that directed forgetting where participants are instructed to forget items does not affect words which are encoded distinctively (e.g. read aloud), this therefore suggests that directed forgetting is restricted to memories which are weaker, such as items which are read silently.

Further research conducted by Castel, Rhodes and Friedman (2013) studied the effect of predicting memory where participants made judgements of their learning of the production effect to discover individuals' awareness of distinctive cues which can improve memory. They found that participants produced higher judgements for learning where participants predicted they would be able to recall information later for produced items than silently read items. Castel, Rhodes and Friedman (2013) stated that producing items by saying them aloud gives specific access to memory and also improves metacognition.

There has been a very limited amount of research investigating the effect that vocalisation of names has on an individual's ability to recall names to faces. Pariante's (1990) study involved participants being presented with images of people's faces with the corresponding name appearing underneath each face. At the presentation phase, where participants were initially shown the names and faces, participants vocalised, mouthed, read silently or heard each name. Performance was determined by the number of correct name-face pairings accumulated in the cued recall stage. Unlike previous research suggesting that vocalisation improves recall for word list stimuli (MacLeod, Gopie, Hourihan, Neary and Ozubko (2010) and

Rackie, Brandt and Eysenk (2014)), participants who read each name silently performed better overall than participants who vocalised names.

Pariante (1990) claimed that this was due to the auditory characteristic of vocalisation (when a name is spoken aloud it is also heard), and this hindered the participants' performance in the vocalisation condition. In contrast, other researchers such as Arenberg (1977) and Greenlee et al (2011) have claimed that the auditory effect of vocalisation is beneficial for individual's memory of items. The contrast in results presented here could be due to the differences in samples, Pariatne's (1990) sample consisted of participants who were all above the age of 60, because memory declines with age (Tong et al, 2013), older peoples' memories are inferior to the memories of younger individuals. The results from Pariante's (1990) study therefore may not be generalisable to the wider population such as university students who are much younger.

Rationale for current research

This research will investigate the effect of testing on name-face paired learning. The testing effect has been investigated by many researchers such as Szpunar et al (2008) and Carpenter (2009), however these studies have largely investigated the recall of word stimuli. Although there has been extensive research examining the testing effect in word list learning domain, there is limited research investigating the utility of this effect in name-face paired learning. Additionally, there has been very little research investigating why name-face pairings become more difficult to learn over time (e.g. as an individual learns more pairings their performance deteriorates) if they are in a large social group or are required to learn a large number of names for their work.

Many studies conducted on testing and recall such as Szpunar et al (2008) and Weinstein et al (2011) have suggested that testing during the presentation of information lessens the build-up of proactive interference and hence participants perform better when they are tested than when they are untested. Much of the literature has demonstrated that testing improves retention of information using a cumulative test where participants are tested between the presentations of information and then complete a test on all of the information they have seen. Weinstein et al (2011) investigated the affect that proactive interference has on name-face recall. Participants were presented with 12 trials (where name-face pairings were presented in combination) in one list and four lists in total. Half of participants were in the tested condition, where they completed a cued recall test after the presentation of each list. The other half of participants were in the untested condition and were only tested on the information they had studied at list four. The correct number of name-face pairings was compared between participants in the tested and untested conditions on the test at list four and on a final cumulative test where participants were required to state the name of each face from all of the lists. The authors in this study concluded that the participants in the tested condition performed significantly better than participants in the untested condition on the test at list four. It could therefore be concluded that testing between the presentations of stimuli minimised the build-up of proactive interference for participants who were in the tested condition as they were able to recall a greater number of name-face pairings. It could be concluded that the information learned from lists 1-3 disrupted the learning of the fourth list more so for the untested participants who remembered fewer name -face pairings than the tested participants. This study also found that participants who were tested between lists performed approximately four times better on the final cumulative test than participants who were tested at list four only (participants in the untested condition).

To ensure that the potential impact of proactive interference is addressed in the present study, participants in each condition will be required to learn four lists of name-face pairings (similarly to the study conducted by Weinstein et al (2011)) and will all be given a cued recall test on the content in the fourth list as well as a cumulative test where they will be required to name each face across all of the lists. It could be suggested that participants who performed better at the test on the fourth list are less affected by proactive interference, and hence their learning of name-face stimuli in prior lists did not inhibit the learning of the stimuli in the fourth list.

Although researchers such as Carpenter and DeLosh (2005) have investigated the difference between the performance recall of participants who have been tested between lists of stimuli and who have restudied stimuli, there has been a limited amount of research using name-face pairings as stimuli and it is important that this is addressed. There were two controlled (i.e. non-confounded) experiments identified in the literature review which investigated how testing affects memory for name-face pairings; the experiment by Carpenter and DeLosh (2005) and the experiment by Weinstein et al (2011). However, these two experiments (as well as the other ones discussed) have been inconsistent in the type of control condition employed and this is the major issue which needs to be addressed when examining the effect of testing on memory for name-face pairs.

In the word list literature Szpunar et al (2008) identified and addressed a key confound in the testing literature - the improved memory retention attributed to the effect of testing could be due to additional exposure to the material during the test rather than the act of retrieval benefitting memory. In order to address this they highlighted the importance of using restudy as a control condition. However, they also acknowledged that restudy itself cannot act as a true control condition as participants are exposed to all the study materials whereas tested participants are only exposed to material they can recall (in the word list literature). This means that restudy has an additional advantage that testing does not and the authors discuss how restudy is biased against testing due to additional and correct exposure, the implications of which is that if a testing effect is demonstrated then it is even more powerful than restudy which is biased against the effect of testing. Following on from their experiment controlling for this in the word list literature it was demonstrated that testing was superior to restudy.

In the name-face paired learning literature restudy differs from the word list literature. Participants are re-exposed to all material in the restudy condition (face as well as associated name) but exposed to the cue (face) in the tested condition (as it is a paired associate task of name-face pairs) and so in the tested condition they experience re-exposure to the cue to which they need to correctly produce the associated name for. Therefore the partial exposure in the name-face paired learning literature means findings cannot be extrapolated from the word list learning literature to establish whether testing is superior to restudy. Consistent with the word list literature this control is biased towards restudy which allows re-exposure to the paired information and against testing which only provides the cue for the pair to be remembered.

The other experiment in the field of name-face pairing (Carpenter & DeLosh, 2005) compared the effect of testing to restudy rather than a distractor task and demonstrated that testing was superior to restudy. However, such studies ignore the potential improvement of recall that participants experience by seeing the material twice, restudy is a learning strategy and should also be investigated in its own right. In the present experiment as well as including a tested condition and a restudy condition, a further, untested condition where participants have no further learning opportunities will also be incorporated. Therefore, as well as comparing the performance of participants in the tested condition to those in the restudy and untested conditions, the restudy condition will also be compared to the untested condition. This experimental design can address the methodological issues outlined above as well as providing insight into the utility of restudy being used as a learning strategy in the field of name-face paired learning.

Considering the research presented above, the present study aims to investigate the effect of testing where participants are either tested between lists or restudy material from lists 1-3 and are tested at list four only which lead to the formation of the following hypotheses:

1. Participants who are tested between the presentations of each list of stimuli (tested) will recall significantly more name-face pairings than participants who restudy material on lists 1-3 (restudy) on a test of the fourth list.
2. Participants who are tested between the presentations of each list of stimuli (tested) will recall significantly more name-face pairings than participants who restudy material on lists 1-3 (restudy) on a cumulative test.

Methodology and project design experiment 1

Design

This was an independent measures experiment; each participant was randomly assigned to one of three conditions. The independent variable was testing which consisted of three conditions; the restudy condition, the tested condition and the untested condition. Two dependent variables were recorded: correct recall and intrusion rates (Weinstein et al, 2011).

The mean number of correctly recalled name-face pairings was compared between the conditions for the test at list four and at the final cumulative test. Data at the test on list four was also obtained for the number of prior list intrusions (names given at list four which are from a different list), current list intrusions (names from within the same list which are paired with the wrong face), extra list intrusions (names which have not appeared in any of the lists) and omissions (where no name has been provided) and these were also compared between the three conditions.

Participants

Participants were 48 undergraduate students at the University of Huddersfield (16 participants per condition), they were recruited to participate in the study via the online university study recruitment system (sona) and were assigned course credit for participating in the study.

Apparatus and stimuli

Adobe Flash Player was used to run the experiment where participants aimed to remember a series of name-face pairings. 48 photographs of male faces were presented to participants and were taken from the neutral set of photographs at the Psychological Image Collection at Stirling database (http://pics.psych.stir.ac.uk/). These photographs were the same as those used in Weinstein et al's (2011) study. One of 48 popular male first names (see Appendix) all of which had two syllables and were between 5-7 letters long were randomly assigned to each photograph and appeared underneath each image in size 28 Arial font when they were presented to participants.

Procedure

Participants were seated directly opposite a computer screen located in a lab room for the duration of the experiment. The experiment consisted of two phases, the first was the presentation phase where participants were shown 48 name-face pairings which consisted of 4 lists (12 name-face pairings in each list) which they were instructed that they should aim to remember - full lists were presented to participants in a random order. The second phase was where each condition differed, a third of participants were given a cued recall test after the presentation of each list where participants aimed to remember the names of the faces they had seen in the immediately preceding list (tested condition). A third of participants were shown the same list again (restudy condition) and a third of participants completed maths questions after each list and were tested on the fourth list only (untested condition).

Instructions

Participants were informed that they were required to study four lists consisting of 12 name-face pairings and that after each list was presented they may or may not complete math problems, may or may not be presented with the same list of name-face pairings again or may or may not be required to attempt to name the faces from the immediately preceding list in a cued recall test. They were told that after the completion of the fourth list they would complete a cumulative test where they would be presented with each face which they would then have to name. During the presentation of the lists participants were told that what follows the presentation of each list is determined randomly by the computer program. In actual fact, there were only three testing schedules; in the tested condition participants had a 60 second (s) maths test followed by a cued recall test lasting 100s after each list. The participants in the restudy condition completed 60s of maths after lists 1-3 and were then presented with lists 1-3 a second time for 100s, after the maths test on the fourth list they completed a cued recall task on list four for 100s. In the untested condition participants had a maths test lasting 160s at lists 1, 2 and 3 and 60s of maths problems as well as a 100s cued recall test after the fourth presented list (Weinstein et al, 2011). Math problems are a combination of simple multiplication, divisions, additions and subtractions and participants were told to complete them as quickly and accurately as possible. This was a distractor task and designed to prevent the rehearsal of previously learned material.

There was a 2s inter stimulus interval between each list and the math test, as well as a 2s inter stimulus interval between each math test and the cued recall phase. At the presentation phase, name-face pairs were presented for 4s each; with a 0.5s inter-stimulus interval between the presentations of each name-face pairing (Weinstein et al, 2011). When participants in the restudy condition were presented with the name-face pairings a second time, they saw them for 8s each. In the initial cued recall tests (which occur after each list), participants were tested on the content of the immediately preceding list. Each face was presented in a random order for 8s and participants were required to type in the name which the face was paired with at the presentation phase. Participants were told not to guess the name if they could not remember it.

After participants completed the cued recall test at the fourth list they were required to complete a final cumulative cued recall test, where they recalled names to all of the 48 faces which they had seen over the four lists. On this test, all of the faces were presented in a random order and participants attempted to recall the name that had been paired with each face within the four lists at the presentation phase. This was completed in the participants' own time and they were told not to guess the name if they could not remember it. They instead pressed a 'skip' button which resulted in the presentation of the next face. If participants could remember the name they typed it underneath the picture of the face; consistent with how participants responded when tested after the lists.

Scoring

For the test at list four, if it was clear participants were in the middle of typing a name and omitted the last couple of letters because they ran out of time to complete the name during the 8s opportunity for retrieval, their responses were marked as being correct. This was however not necessary at the cumulative test as participants had as much time as they required to complete each trial. Participant's responses were correct if the name they provided was exactly the same allocated name which matched the face, names which were provided that were close but had a misspelling were also accepted as being correct. For example, ‘Steven’ was accepted in place of ‘Stephen’ and ‘Aarron’ was accepted in place of ‘Aaron’.

Results for experiment 1

Where there are significant differences, post hoc tests will be conducted to discover where the differences lie; the alpha level will be at .017 to account for multiple comparisons. Cohen’s d will be analysed to assess the effect size of testing on the dependent variable.

Table 1 Means and standard deviations (2dp) (in brackets) of correct responses, prior list intrusions, current list intrusions, extra list intrusions and omissions for participants in the tested condition when tested on lists 1-4.

Response	List 1	List 2	List 3	List 4
Correct	6.00 (2.10)	5.00 (2.03)	4.4 (2.07)	4.4(2.71)
Prior list intrusion	-	.69 (.80)	1.13 (1.20)	1.13 (.89)
Current list intrusion	2.31 (1.50)	1.94 (1.44)	1.94 (.93)	1.38 (1.03)
Extra list intrusion	.88 (.89)	.94 (1.20)	1.13 (1.36)	1.19 (1.05)
omission	2.81 (2.17)	3.44 (1.87)	3.44 (2.28)	3.94 (2.27)

Table 1 shows the number of correctly recalled name-face pairings reduces from each list to the next. The number of prior list intrusions rises from lists 2 to 3 and then remains constant at list four. The number of current list intrusions decreases from each list to the next while the number of extra list intrusions increases from each list to the next. The number of omissions also increases from each list to the next.

List four test

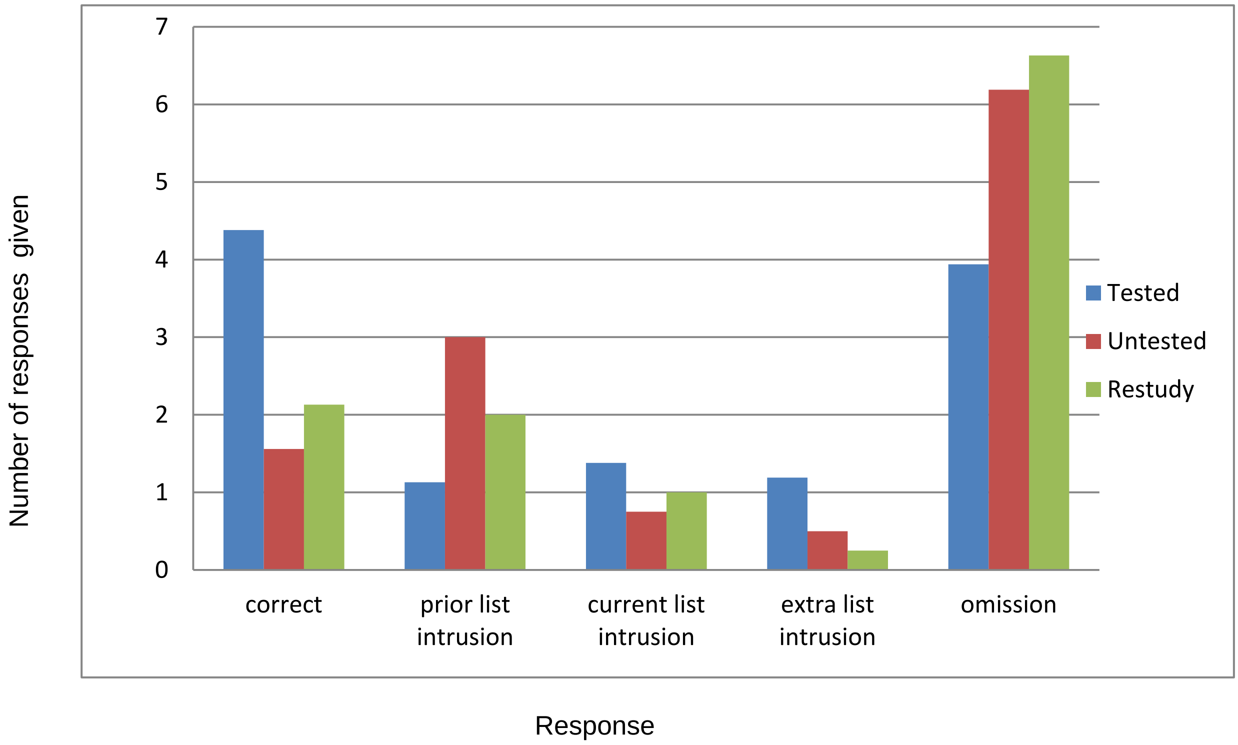

Figure 1 The mean number of correct responses, prior list intrusions, current list intrusions, extra list intrusions and omissions for all conditions when tested at the fourth presented list.

Figure 1 shows that participants who were tested between the presentations of each list of name-face pairings performed better than participants who were untested and participants who restudied the stimuli. The number of prior list intrusions is higher for participants in the untested condition than participants in the other conditions. The number of current list intrusions appears to be similar in all three conditions while tested participants have a higher number of mean extra list intrusions. Compared to the other response categories, the number of omissions is high for every condition, however tested participants have a lower mean number of omissions than participants in the untested and restudy conditions, this is because their responses is higher for some of the other categories.

A Shapiro-Wilk test of normality was conducted and this suggested that the data from all conditions were normally distributed ($p > .05$). A Levene's Test of homogeneity of variance confirmed that variances were statistically equivalent ($p > .05$).

A one way ANOVA was used to analyse whether the independent variable of testing significantly affected the performance of participants. There was a main effect of testing, ($F(2,45) = 10.129$, $p < .001$) there was a significant difference between the number of correct responses provided by the participants in each condition. A Tukey post hoc test was conducted and this suggested that participants in the tested condition provided a significantly higher number of correct responses than participants who were in the untested condition ($p < .001$, $d = 1.40$) and participants who were in the restudy condition ($p = .004$, $d = 1.10$). Both Cohen's d effect sizes suggest that testing had a large effect on participants' ability to recall correct name-face pairings at the test of list four, therefore testing results in an increase of correct responses at list four when compared to both the untested and restudy conditions. There was no significant difference between the number of correct responses in the untested and restudy conditions ($p = .674$, $d = .70$). However the Cohen's d effect size indicates that there was a medium effect size on recall at list four between untested and restudy participants where restudy participants recalled more correct name-face pairings than untested participants at the test on list four.

A one way ANOVA also suggested that there was a significant difference between the number of prior list intrusions provided by participants across the three conditions ($F(2,45) = 5.205$, $p = .009$). A Tukey post hoc test suggested that participants in the tested condition provided significantly fewer prior list intrusions on the test at list four than participants in the untested condition ($p = .007$, $d = 1.20$) but not for participants in the restudy condition ($p = .298$, $d = .61$). The Cohen's d effect size suggests that testing had a large effect on the number of prior list intrusions provided when compared with the untested condition and a medium effect size when compared with the restudy condition, therefore testing decreases the number of prior list intrusions given by participants. There was no significant difference between the number of prior list intrusions provided between participants in the untested and restudy conditions ($p = .209$, $d = .52$). The Cohen's d effect size suggests that there was a medium size effect on the number of prior list intrusions given between untested and restudy participants and restudy participants gave less prior list intrusions than untested participants.

There was no significant difference between the number of current list intrusions in each condition ($F(2,45) = 1.748$, $p = .186$).

A one way ANOVA was used which suggested that there was a significant difference between the number of extra list intrusions in each condition ($F(2,45) = 6.185$, $p = .004$). A Tukey post hoc test indicated that participants in the tested condition provided significantly more extra list intrusions than participants in the restudy condition ($p = .004$, $d = 1.17$) but not for participants in the untested condition ($p = .043$, $d = .77$). The Cohen's d effect size is large for testing when compared to the restudy condition and was medium when compared to the untested condition, where tested participants gave more prior list intrusions than restudy and untested participants. There was no significant difference between the number of extra list intrusions given between the untested and restudy conditions ($p = .640$, $d = .41$). The Cohen’s d effect size shows that there was a small effect size on the number of provided extra list intrusions, restudy participants gave fewer than untested participants.

A one way ANOVA suggested that there was a significant difference between the number of omissions provided for participants in each condition ($F(2,45) = 7.159$, $p = .002$). A Tukey post hoc test indicated that participants in the tested condition provided significantly fewer omissions than participants in the untested condition ($p = .014$, $d = 1.11$) and restudy conditions ($p = .003$, $d = 1.16$). Both Cohen’s d effect sizes suggest that testing has a large effect size when tested participants are compared with untested and restudy participants, tested participants gave fewer omissions than untested and restudy participants. There was no significant difference between the number of omissions given between participants in the untested and restudy conditions ($p = .835$, $d = .20$). The Cohen’s d effect size suggests that there was a small effect size on the number of omissions given, restudy participants gave slightly more omissions than untested participants.

Cumulative test

Table 2 The mean number, standard deviations (SD) and confidence intervals of correctly recalled name-face pairings for participants in all conditions at the cumulative test.

Condition	N	Mean (2dp)	SD (2dp)	95% Confidence Interval for mean	
				Lower	Upper
Tested	16	13.13	8.48	8.60	17.65
Untested	16	3.31	1.70	2.41	4.22
Restudy	16	7.81	4.15	5.60	10.02

Table 2 indicates that participants who were tested after the presentation of each list of name-face pairings performed better on a cumulative test than participants who were not tested after each list and participants who restudied the information. Participants in the untested condition have the lowest mean number of recalled pairings from the three conditions. The standard deviation for participants in the tested condition is higher than the other conditions which suggests that there was more variance in the number of correctly recalled name-face pairings in this condition than in the untested and restudy conditions.

A Shapiro-Wilk test of normality was conducted and this suggested that the data from all conditions were normally distributed ($p > .05$). A Levene's Test of homogeneity of variance confirmed that variances were statistically equivalent ($p > .05$).

A One Way ANOVA was conducted to discover whether there was a significant difference between participants who were in the tested, untested or restudied condition on the number of correctly recalled name-face pairings on the cumulative test. This suggested that there was a main effect of testing ($F(2,45) = 12.573$, $p < .001$). A Tukey post hoc test was conducted and this suggested that participants in the tested condition recalled significantly more correct name-face pairings on the cumulative test when compared to participants in the untested condition ($p < .001$, $d = 1.54$) but not the restudy condition ($p = .025$, $d = .80$). Both Cohen's d effect sizes suggest that testing had a large effect on participants' ability to recall correct name-face pairings on the cumulative test, therefore testing results in an increase of correct responses when compared to both the untested and restudy conditions. There was no significant difference between the number of correctly recalled name-face pairings at the cumulative test between participants in the untested and restudy conditions ($p = .066$, $d = 1.42$). The Cohen's d effect size suggests that there was a large effect on recall at the cumulative test between untested and restudy participants and restudy participants recalled more correct name-face pairings than untested participants at the cumulative test.

Discussion for experiment 1

The results suggested that there was a main effect of testing between the presented lists of name-face pairings on the test at list four; participants in the tested condition recalled significantly more name-face pairings than participants in both the untested and restudy condition at the test at list four, the first hypothesis was therefore accepted.

The results suggested that participants in the tested condition recalled significantly more correct name-face pairings than participants in the untested condition at the cumulative test where they were tested on all of the name-face pairings they had seen across the four lists. Although tested participants performed better than restudy participants at the cumulative test there was no significant main effect of testing when the alpha level had accounted for multiple comparisons; the second hypothesis was therefore rejected.

One limitation of the present experiment was its relatively small sample size, as effect sizes are not affected by sample size they were investigated. The Cohen's d effect sizes show that testing between the presentations of name-face stimuli has a very large effect on the correct recall of name-face pairings when compared to untested and restudy conditions at both the list four test and at the final cumulative test. Although the difference in performance determined by the number of correctly recalled name-face pairings was not significant at the cumulative test between tested and restudy participants the effect size was large which suggests that testing does assist individuals with retention of information when compared to restudying material.

It would therefore be beneficial for individuals to test themselves when they are in a situation when it would be advantageous to learn many individuals' names.

The effect sizes were large when comparing testing and restudy conditions for both the test at list four and the cumulative test. This suggests that testing helps individuals to remember name-face pairings more so than restudying information when they are presented with a few faces and when they are required to learn a high number of name-face pairings such as the 48 which were presented to participants in the present study. There was no significant difference between restudy and untested conditions at both the list four test and at the cumulative test. There was however, a medium effect size when comparing the two conditions at the test on list four and a large effect size when comparing the two conditions at the cumulative test. This suggests that although there was no significant difference in results, restudying material may assist individuals to recall names more so than not having any additional learning opportunities, particularly when an individual has lots of names they are required to remember.

The standard deviations at the cumulative test are large for both the tested and restudy conditions. This suggests that results were vastly spread out from the mean and the number of correctly recalled name-face pairings may have overlapped from each condition. This indicates large individual differences between those who benefit from restudy and those who benefit from the effect of testing. Further research could therefore investigate the boundaries of the testing effect to identify who will benefit most from this strategy.

The results obtained in the present experiment are similar to the studies conducted by Nunes and Weinstein (2012), Szpunar et al (2008) and Weinstein et al (2011) which found that testing between lists of stimuli improved the performance of recall when tested on a final list and on a cumulative test. At the cued recall test on list four participants who were tested performed significantly better than participants who were in the untested and restudy conditions, it could therefore be argued that the effect of proactive interference was reduced in tested participants. It could be suggested that information which had been previously learned (from earlier lists) negatively affected participants' ability to recall name-face pairings from further lists particularly for participants who were in the untested and restudy conditions.

The number of prior list intrusions (where names have been reported which were presented in earlier lists) can be used to explain how proactive interference has affected results. Proactive interference has a greater effect on participants who have a larger number of prior list intrusions as this shows that they have some difficulty differentiating between the list they are being tested on at the current time and the lists that they have previously seen. In the study conducted by Szpunar et al (2008), participants in the tested group recalled twice as many words in the free recall test on the fifth list than participants in the untested group. This research found that those who were tested on lists 1-4 produced a significantly lower number of prior list intrusions (words which were presented in lists 1-4) at list 5. However, when participants in the untested condition were required to recall words from the list five test, they produced as many prior list intrusions as they did correct responses. Similarly, in the Weinstein et al (2011) paper, there were hardly any prior list intrusions for participants in the tested condition whereas participants in the untested condition recalled a similar number of correct responses (mean = 2.88) to prior list intrusions (mean = 2.44). The findings from the present experiment were similar in regard to the number of prior list intrusions; those in the untested condition provided significantly more prior list intrusions than participants in the tested condition.

It could be argued that the results of the present experiment and the findings obtained from Szpunar et al (2008) and Weinstein et al (2011) support the notion that when participants are given the opportunity for retrieval between the presentation of information they reset their encoding processes (Pastötter et al, 2011). This is the reason why the effect of proactive interference is reduced for participants who were tested between the presentation of stimuli. Untested participants provided more prior list intrusions which suggest that their encoding processes were not reset as in the final test list they recalled information from previously presented lists. Participants who were tested between the presentations of stimuli, however, were better able to distinguish between the materials in each list and therefore performed better on the test at the fourth list. However, the mean number of correctly recalled name-face pairings in the tested condition decreased from lists 1 to 4, so although encoding processes were reset, it could be more effective to learn one list rather than four. Although there was no significant difference between the number of prior list intrusions between participants in the tested and restudy condition, it appears that proactive interference had a greater impact on results for participants in the untested and restudy conditions, considering they performed significantly worse than those in the tested condition (identified from recalling significantly less correct name-face pairings). Proactive interference must also have affected results in the tested condition to a certain extent as the number of correct name-face pairings reduced from lists one to four and the number of prior list intrusions from lists two to four increased.

There were more current list intrusions (names from within the same list which are paired with the wrong face) and extra list intrusions (names which have not appeared in any of the lists) present than prior list intrusions for participants in the tested condition on the test at list four. It could therefore be argued that although testing assists individuals with list segregation (identifying which faces and names came from a certain list) but it does not assist individuals with within list segregation i.e. participants are able to remember the names of those within the previously presented list but not necessarily the correct name for each face within a certain list.

The number of omissions (where no name had been provided) rises from list to list in the tested condition; this suggests that participants remembered less name-face pairings as the lists progressed. Participants were instructed not to guess if they could not remember a name, the more lists of name-face pairings participants learned the more they were unable to provide a response. Tested participants provided significantly fewer omissions at the list four test when compared to untested and restudy participants. This suggests that tested participants attempted to recall more name-face pairings at the list four test.

Participants in the tested condition performed significantly better than participants in both the untested and restudy condition at a final list test and they also performed significantly better than untested participants at a cumulative test. Although tested participants recalled more correct name-face pairings when compared to restudy participants at the cumulative test, the result was not significant. This finding is supported by other research which obtained similar data such as Szpunar et al (2008) experiment. The results from the present study are similar to those obtained by Szpunar et al (2008) in regard to the performance of participants in the restudy and tested conditions. Tested participants performed significantly better than participants who restudied information, these results further suggest that it is the process of being tested, rather than being re-exposed to stimuli that improves performance on a cued recall task.

Carpenter and DeLosh (2005) identified results which are consistent with the present experiment's findings that tested participants perform better on a cumulative cued recall test than participants who restudy material. Carpenter and DeLosh's (2005) experiment used the same number of stimuli as the present experiment (48 faces). Although this study split the stimuli into lists the performance of each list was not assessed, they did not therefore account for the number of recalled name-face pairings at different stages of the presentation phase and how the number of recalled items varied from list to list. The potential effect of proactive interference was therefore not investigated. The number of intrusions provided from each list could also not be obtained, therefore the authors of the study would be unable to determine the extent to which testing aided participants other than improving the recall of name-face pairings, for example investigating prior list intrusions is important to understand that encoding processes had been reset from one list to another. The study conducted by Carpenter and DeLosh (2005) used restudy as a control condition and therefore ignored the notion that restudy in itself is a learning strategy.

The present study, however, used a further condition (untested condition) where participants had no further learning opportunities as a control. In the present experiment, there was no significant difference between the performance of participants in the untested condition and the performance of participants in the restudy condition across all of the dependent variables including intrusions rates. There was, however, a medium size effect between untested and restudy participants at test four and a large effect size between untested and restudy participants at the cumulative test, where restudy participants recalled more correct name-face pairings than untested participants. The results from the present experiment therefore identified the usefulness of restudying material as a strategy in itself.

In the present experiment, although participants in the tested condition did perform better than participants in the untested and restudy condition, they did not recall as many correct pairings (in both the tested and untested condition) as participants in Weinstein et al's (2011) experiment. In comparison to Weinstein et al's (2011) paper the present experiment found that participants in the tested condition provided significantly fewer omissions on the test at list four than participants in the other conditions, although the number of omissions was high across all conditions. This could be because in the present study tested participants recalled more names and therefore attempted to respond more at the testing phase and therefore made fewer omissions. Results for intrusions were however generally consistent between Weinstein et al (2011) paper and the present experiment.

The findings from the present experiment are consistent with much of the previous literature (eg, Weinstein et al (2011) & Helder & Shaughnessy (2008)) which has outlined that testing at intervals throughout the presentation of name-face pairings increases the retention of to be remembered information. This therefore suggests that the present research has high external validity. There have been many explanations which have been used to explain why individuals who are tested generally perform better than those who are not tested on a cued recall task. Szpunar et al (2008) stated that testing between lists of information leads to better discrimination of stimuli; participants are therefore able to remember information which was learned in a particular list. This explanation of the testing effect argued that testing an individual on the content of each list of information meant that when individuals are required to retrieve information they are able to restrict the search of material to the most recently learned list. This is done by separating information which has been previously tested from information which has been learned but has yet to be tested. The findings from the present experiment support this explanation, tested participants made fewer prior list intrusions but higher current and extra list intrusions so this demonstrates that testing increases discrimination between lists but not within lists.

Kuo and Hirshman (1997) outlined that testing an individual's memory differentiates items from other items which have been learned in previous encoding episodes and therefore learned in previous lists. This then generates the processing of item specific features of the items in the tested lists (Carpenter & DeLosh, 2006). Item specific information which enables specific target items to be identified within a search and relational information which organises the memory search are depended on when there is an opportunity to retrieve information (through testing) (Matthews, Smith, Hunt & Pivetta, 1999). This could serve as an explanation of why stimuli which is tested at intervals throughout the presentation of stimuli is better remembered than stimuli which is untested and restudied. Convergence of item specific and relational information is better when an individual has the opportunity of retrieval compared to when an individual has no opportunity for retrieval such as those who are untested or restudy material (Matthews et al, 1999).

Pastötter, Shicker, Niedernhuber and Bäuml (2011) further investigated the concept of the testing effect. They found that the efficiency of encoding of to be remembered stimuli was maintained when participants were tested between lists of presented material. Pastötter et al (2011) recorded EEGs of participants throughout the encoding stage where participants were required to learn the information. When participants were presented with several lists they found that there was no increase in alpha levels when they were tested which suggests that the attention of participants was maintained. They therefore concluded that testing resets the encoding of information when it is at its maximum efficiency for the learning of material on the next to be studied list.

Experiment 2

It would be beneficial for individuals who are required to learn many name-face pairings - for example in work or social scenarios - to know the optimal way of learning names of people. Therefore, a fundamental aim for cognitive psychologists is to elucidate the factors that can improve memory span and in this domain the typical factors identified consist of mnemonic strategies designed to facilitate memory for the information to be learnt. It is therefore important to establish whether additional variables other than testing can be used to improve performance. Ideally the strategies identified should be able to be used in combination without interfering with each other. A major issue to consider when identifying strategies is whether they would work outside the laboratory.

However, the mnemonic strategies identified in the literature are often very complex and so attempting to apply them without explicit training may not be feasible. For example, one strategy involves noting the person's name and then inspecting their appearance for a distinctive feature followed by constructing a mental image that associates the name to that feature. Although this technique has demonstrated marked improvements under experimental conditions (McCarty, 1980; Morris, Jones, and Hampson, 1978) and is often promoted in the memory improvement literature (e.g., Gruneberg and Herrmann, 1997; Higbee, 2001) the technique has failed to demonstrate any benefit when implemented in more ecologically valid experiments. When asked to use this mnemonic strategy whilst maintaining a conversation it does not improve recall of names (Patton, 1994). The technique places too much demand on cognitive resources when used in situations where resources are already diverted to maintaining an ongoing conversation. Therefore the importance of identifying more naturalistically applied techniques remains a pertinent issue. Additionally, Carpenter and DeLosh (2005) investigated spacing effects as a further variable, as well as testing. However, although testing can be controlled for by the individual, spacing effects cannot as an individual cannot control how many people they come into contact with in a certain amount of time. This technique therefore is unlikely to aid the recall of name-face pairs in a real life setting. Therefore as well as testing, it is important to investigate other variables which can be controlled for by the individual.

What remains to be established is whether there are other variables which could conceivably improve memory for name-face pairs in addition to that of testing. The purpose of the second experiment is to identify further variables which could be implemented alongside testing to improve memory for name-face pairs. The importance of identifying further variables would be of substantial benefit in certain fields which emphasise face and name recognition. One candidate variable which complements learning of names and can be used in conjunction with testing is vocalisation. The empirical literature indicates a clear benefit of reading words (such as names) aloud. The production effect is well established in the cognitive psychology literature whereby memory for words read aloud (i.e., produced) is much stronger than memory for words read silently. MacLeod et al (2010) argued that the phenomena results from enhanced distinctiveness: words read aloud have supplementary additional unique information that is beneficial at test for discriminating produced words from other words.

In addition to it being one of the simplest mnemonic techniques identified in the literature vocalising names is a naturalistic strategy used by individuals when meeting new people where the person listening may repeat someone's name to confirm it or vocalise the name if they refer to the person they have been introduced to in later conversation (Patton, 1994). However, there has been little research directly addressing how vocalisation can improve memory for name-face pairs when participants are young and healthy and no research that the authors are aware of which investigates the combined effect of testing and vocalisation on memory for name-face pairs. It has been demonstrated that young children have a superior memory for object labels when they vocalise compared to non-vocalisation groups, indicating that vocalisation can be used as a mnemonic tool when remembering labels for objects (Icht and Mama, 2015).

However, it is not clear how well this mnemonic strategy works in a paired associate task when the name to be learnt is not meaningful (Carpenter & DeLosh, 2005). In addition to this the production effect has primarily been tested using recognition tasks (e.g., Forin et al, 2012 and Ozubko, Hourihan & MacLeod, 2012) and so it is not clear how useful the technique will be when the person has to recall the name associated with the face from memory rather than recognise the name when it is visually presented to them. It is generally accepted that recognition memory tasks are easier than recall tests of memory and so people typically perform superiorly in the former than the latter (Haist, Shimamura, & Squire, 1992). However, the current task requires recall of a name associated with a unique face. This procedure allows us to stimulate an everyday learning situation where the demand is for a person to recall the name when they see a face rather than having the name visually presented for recall (Weinstein et al, 2011).

Consistent with experiment 1, experiment 2 initially tested the assumption that memory for name-face pairs could be enhanced by testing of stimuli. If this effect of improved memory retention of name-face pairs due to testing was replicated, the novel aim of experiment 2 was to assess whether vocalisation improved memory retention in name-face learning. This study design allows the author to establish if vocalisation can improve memory for name-face paired learning. If vocalisation of names results in the hypothesised improvement in memory for name-face pairs the approach could be adopted to improve learning of face–name pairs in everyday situations (e.g., social or work functions) or in fields where employees are required to keep track of large sets of name-face pairs. Interestingly this is a behaviour people engage in naturally (repeating ones name after hearing it) so a negative effect would be just as informative as a positive one (Patton, 1994).

The current experiment adapted the Weinstein et al (2011) paradigm (originally adapted from Szpunar et al (2008) to assess the effect of testing on word list learning) to assess both the effect of testing on memory for name-face pairs as well as the effect of vocalisation and any additive effects from the combined use of the two strategies. Thus the design of the experiment incorporated a vocalisation and no vocalisation of names group as well as a tested and restudy group. This extension of experimental conditions allowed the researcher to simulate an everyday learning situation while addressing the theoretical question outlined above with regard to strategies for improving name-face learning.

Considering the research presented above, the present study aims to investigate the effect of vocalisation where names will either be read silently or aloud which lead to formation of the following hypotheses:

3. Participants who vocalise names will recall significantly more name-face pairings than participants who read the names silently on the cued recall test at list four.
4. Participants who vocalise names will recall significantly more name-face pairings than participants who read the names silently on the cued recall test at the final cumulative test.

Methodology and project design experiment 2

Design

This was an independent measures experiment; each participant was randomly assigned to one of three conditions. Experiment 2 assesses the impact of two independent variables; testing which had two levels, either testing or restudy and vocalisation which had two levels, vocalised or not vocalised. There were three conditions: tested vocalise, tested not vocalise and restudy. The dependent variables were correct recall and intrusions rates at list four. To assess the effect of vocalising on name-face recall performance was compared between both the tested not vocalise and tested vocalise conditions. The effect of testing was further assessed between the participants in the testing not vocalise condition and the restudy condition. The mean number of correctly recalled name-face pairings was compared between the conditions for the test at list four and at the final cumulative test.

Participants

Participants were 60 undergraduate students at the University of Huddersfield (20 participants per condition). There were 17 female participants and 3 male participants in the tested vocalise condition with a mean age of 22.15. There was also 17 female and 3 male participants in the tested not vocalise condition and the mean age for this condition was 21.85. Additionally there were 18 female and 2 male participants in the restudy condition with a mean age of 22.25. All participants were recruited to participate in the study via the online university study recruitment system (sona) and were assigned course credit for participating in the study.

Apparatus and Stimuli

Experiment 2 will use the same program that was used to run experiment 1 and the same stimuli.

Procedure

Participants were seated directly opposite a computer screen located in a lab room for the duration of the experiment. The experiment consisted of two phases, the first was the presentation phase where participants were shown 48 name-face pairings which consisted of 4 lists (12 name-face pairings in each list) which they were instructed that they should aim to remember - full lists were presented to participants in a random order. A third of participants were in the tested vocalise condition, here participants were required to say each name aloud during the presentation stage. A third of participants were in the tested not vocalise condition and the final third of participants were in the restudy condition. Participants in both the tested not vocalise and restudy conditions were required to read each name silently at the presentation stage. Participants in both the tested vocalise and tested not vocalise conditions were tested on the names of the faces they had learned in each, immediately preceding list, while participants in the restudy condition were shown lists 1-3 twice and then tested at list four.

Instructions

Participants were informed that they were required to study four lists consisting of 12 name-face pairings and that after each list was presented they may or may not complete math problems, may or may not be presented with the same list of name-face pairings again or may or may not be required to attempt to name the faces from the immediately preceding list in a cued recall test. They were told that after the completion of the fourth list they would complete a cumulative test where they would be presented with each face which they would then have to name. Participants in the tested vocalise condition were told to say each name aloud while participants in the tested not vocalise and restudy conditions were instructed to say each name silently. During the presentation of the lists participants were told that what follows the presentation of each list is determined randomly by the computer program. In actual fact, there were only two testing schedules; in both the tested vocalise and tested not vocalise conditions participants had a 60s maths test followed by a cued recall test lasting 100s after each list. The participants in the restudy condition completed 60s of maths after lists 1-3 and were then presented with lists 1-3 a second time for 100s, after the maths test on the fourth list they completed a cued recall task on list four for 100s (Weinstein et al, 2011). Math problems are a combination of simple multiplication, divisions, additions and subtractions and participants were told to complete them as quickly and accurately as possible. This was a distractor task and designed to prevent the rehearsal of previously learned material.

There was a 2s inter stimulus interval between each list and the math test, as well as a 2s inter stimulus interval between each math test and the cued recall phase. At the presentation phase, name-face pairs were presented for 4s each; with a 0.5s inter-stimulus interval between the presentations of each name-face pairing (Weinstein et al, 2011). When participants in the restudy condition were presented with the name-face pairings a second time, they saw them for 8s each. In the initial cued recall tests (which occur after each list), participants were tested on the content of the immediately preceding list. Each face was presented in a random order for 8s and participants were required to type in the name which the face was paired with at the presentation phase. Participants were told not to guess the name if they could not remember it.

After participants completed the cued recall test at the fourth list they were required to complete a final cumulative cued recall test, where they recalled names to all of the 48 faces which they had seen over the four lists. On this test, all of the faces were presented in a random order and participants attempted to recall the name that had been paired with each face within the four lists at the presentation phase. This was completed in the participants’ own time and they were told not to guess the name if they could not remember it. They instead pressed a ‘skip’ button which resulted in the presentation of the next face. If participants could remember the name they typed it underneath the picture of the face; consistent with how participants responded when tested after the lists.

Scoring

The scoring process for experiment 2 is the same as in experiment 1.

Results for experiment 2

Table 3 Means and standard deviations (2dp) (in brackets) of correct responses, prior list intrusions, current list intrusions, extra list intrusions and omissions for participants in the tested not vocalise condition when tested on lists 1-4.

Response	List 1	List 2	List 3	List 4
Correct	6.05 (2.52)	5.25 (2.27)	4.65 (2.37)	4.35 (2.48)
Prior list intrusion	-	.35 (.59)	1.10 (1.02)	.80 (1.44)
Current list intrusion	2.20 (1.51)	1.80 (1.36)	1.30 (1.55)	1.05 (1.15)
Extra list intrusion	.65 (.81)	.50 (.76)	.85 (1.18)	.60 (1.10)
Omission	3.10 (1.76)	4.10 (2.02)	4.10 (2.02)	5.20 (2.31)

Table 3 suggests that the number of correct responses declines from the learning of one list to the next for participants in the tested not vocalise condition, as do the number of current list intrusions. The number of extra list intrusions is fairly constant across all of the lists while the number of prior list intrusions rises from the second list to the third. Furthermore, the number of omissions increases from each list which suggests that participants failed to provide a response when they were exposed to more material.

Table 4 Means and standard deviations (2dp) (in brackets) of correct responses, prior list intrusions, current list intrusions, extra list intrusions and omissions for participants in the tested vocalise condition when tested on lists 1-4.

Response	List 1	List 2	List 3	List 4
Correct	6.65 (3.05)	5.70 (2.87)	4.95 (2.91)	4.55 (2.61)
Prior list intrusion	-	.55 (.76)	1.75 (2.43)	1.05 (1.64)
Current list intrusion	1.90 (2.23)	1.50 (1.57)	1.15 (1.42)	1.70 (1.72)
Extra list intrusion	.50 (.76)	.95 (1.73)	.50 (.89)	.65 (.93)
Omission	2.95 (3.02)	3.30 (2.56)	3.65 (2.76)	4.05 (2.59)

Table 4 shows, similarly to the participants in the tested not vocalise condition that participants in the tested vocalise condition's performance of correct responses worsens from each list of learned name-face pairs to the next. The number of extra list intrusions is fairly constant and the number of omissions rises from the presentation of each list to the next. Therefore vocalisation does not appear to be either increasing correct performance or reducing intrusions.

List four test

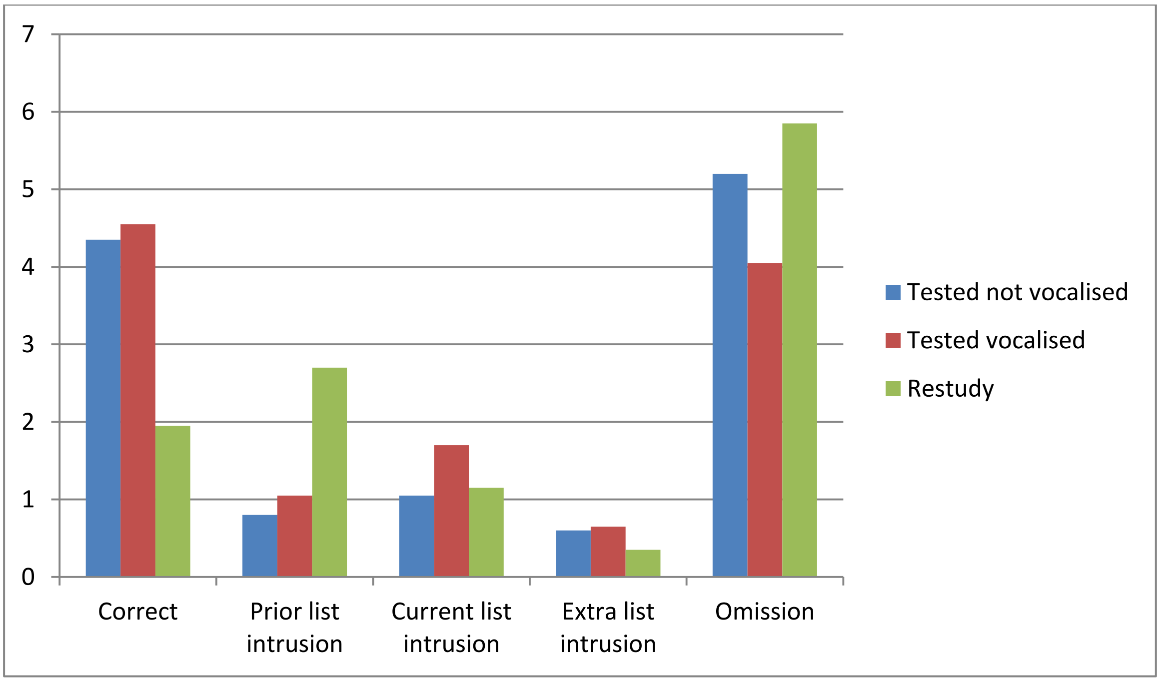

Figure 2 The mean number of correct responses, prior list intrusions, current list

intrusions, extra list intrusions and omissions for all conditions when tested at the fourth presented list.

Figure 2 shows that the number of correctly recalled name-face pairings was much lower for participants who were in the restudy condition than participants who were in both the tested vocalise and tested not vocalise conditions. Additionally, the number of prior list intrusions made by participants in the restudy condition is more than double the number of prior list intrusions from participants in both of the tested conditions. There was more current list intrusions made by participants in the tested vocalise condition than participants in the tested not vocalise and restudy conditions. The fewest number of intrusions were the extra list intrusions of which there were a similar number in all conditions. The number of omissions made was high for every condition although participants in the tested vocalise condition made fewer omissions than participants in other conditions.

A Shapiro-Wilk test of normality was conducted and this suggested that the data from all conditions were normally distributed ($p > .05$). A Levene's Test of homogeneity of variance confirmed that variances were statistically equivalent ($p > .05$).

The effect of testing at list four

A one way ANOVA was used to determine whether there was a significant effect of testing by comparing the restudy and tested not vocalise conditions for performance of participants at list four for correct responses, intrusions and omissions.

There was a main effect of testing at list four ($F(1,38) = 13.067$, $p = .001$, $d = 1.14$), there was significantly more correct responses given by participants in the tested not vocalise condition than participants in the restudy condition. The Cohen's d effect size suggests that testing had a large effect on participants' ability to recall correct name-face pairings at the test of list four, therefore testing results in an increase of correct responses at list four.

A one way ANOVA also suggested that there was a significant difference between the number of prior list intrusions provided for participants in the restudy and tested not vocalise conditions ($F(1,38) = 13.015$, $p = .001$, $d = 1.14$). Participants in the restudy condition gave significantly more prior list intrusions than participants in the tested vocalise condition. The Cohen's d effect size suggests that testing had a large effect on the performance of prior list intrusions given, therefore testing results in a fewer number of prior list intrusions.

A one way ANOVA determined that there was no significant difference between the number of current list intrusions given in the restudy and tested not vocalise conditions ($F(1,38) = .074$, $p = .787$, $d = .09$.). The Cohen's d effect size suggests that testing had a minute effect on the number of current list intrusions given.

Additionally, there was no significant difference between the number of extra list intrusions provided by participants in the tested not vocalise and restudy conditions ($F(1,38) = .604$, $p = .442$, $d = .25$). The Cohen's d effect size suggests that testing therefore had a small effect on the number of extra list intrusions given with tested participants providing more extra list intrusions that restudy participants.

No significant difference was found between the number of omissions for both conditions ($F(1,38) = .765$, $p = .387$, $d = .28$). The Cohen's d effect size suggests that testing had a small size effect on the number of omissions provided at list four, testing therefore only had a slight effect on the number of omissions provided, with tested not vocalise participants providing fewer omissions than restudy participants.

The effect of vocalisation at list four

A one way ANOVA was used to determine whether there was a significant effect of vocalisation by comparing the tested not vocalise and tested vocalise conditions for performance of participants at list four for correct responses, intrusions and omissions.

A one way ANOVA suggested that there was no significant difference between the number of correctly recalled name-face pairings at list four between participants in the tested not vocalise and tested vocalise conditions ($F(1,38) = .062$, $p = .805$, $d = .079$), this confirmed that there was no main effect of vocalisation. The Cohen's d effect size suggests that vocalisation had a minute effect on participants' ability to recall correct name-face pairings at the list four test; vocalisation did not therefore have an effect on correct recall at list four.

Furthermore, there was no significant difference between the number of prior list intrusions between the two conditions at list four ($F(1,38) = .263$, $p = .611$, $d = .16$). The Cohen's d effect size suggests that vocalisation had a minute effect on the number of provided prior list intrusions, vocalisation did not therefore have an effect the number of prior list intrusions provided.

Additionally, there was no significant difference between the number of current list intrusions for the tested not vocalise and tested vocalise conditions ($F(1,38) = 1.978$, $p = .168$, $d = .44$). The Cohen's d effect size suggests that vocalisation had a small to medium effect on the number of current list intrusions given, participants in the tested vocalised group provided slightly more current list intrusions than participants in the tested not vocalise condition.

There was also no significant difference between the number of extra list intrusions for the tested not vocalise and tested vocalise conditions ($F(1,38) = .024$, $p = .877$, $d = .05$). The Cohen's d effect size indicates that vocalisation had a minute size effect on the number of extra list intrusions provided at the list four test, vocalisation did not therefore affect the number of extra list intrusions provided.

There was no significant difference between the number of omissions provided at the list four test for participants in the tested vocalise and the tested not vocalise condition $F(1,38) = 2.203$, $p = .146$, $d = .047$). The Cohen's d effect size suggests that vocalisation therefore had a minute size effect on results; vocalisation therefore did not have an effect on the number of omissions given.

Cumulative test

Table 5 The mean number, SD and confidence intervals of correctly recalled name-face pairings for participants in all conditions at the cumulative test.

Condition	N	Mean (2dp)	SD (2dp)	95% Confidence interval for mean	
				Lower	Upper
Tested not vocalise	20	13.05	7.39	9.59	16.51
Tested vocalise	20	15.10	8.45	11.14	19.06
Restudy	20	6.75	6.08	3.91	9.59

Table 5 suggests that the number of correctly recalled name-face pairings is much lower for participants in the restudy condition than participants in both of the tested conditions, while the mean number of correct responses is higher for participants in the tested vocalise condition when compared to the tested not vocalise condition. The standard deviation is the highest for participants in the tested vocalise condition which indicates that the variance in performance of correctly recalled names was higher than in the tested not vocalise and restudy conditions.

A Shapiro-Wilk test of normality was conducted and this suggested that the data from all conditions were normally distributed ($p > .05$). A Levene's Test of homogeneity of variance confirmed that variances were statistically equivalent ($p > .05$).

A one way ANOVA was conducted to discover whether there was a significant difference between participants' in the tested not vocalise and restudy conditions performance (determined by the number of correctly recalled name-face pairings at the cumulative test). This indicated that there was a main effect of testing ($F(1,38) = 8.764$, $p = .005$, $d = .93$). Participants in the tested not vocalise condition recalled significantly more name-face pairings than participants in the restudy condition at the cumulative test. The Cohen's d effect size indicates that testing had a large size effect on the performance of participants at the cumulative test, testing therefore resulted in an increase of correct recall at the cumulative test.

A one way ANOVA was also conducted to determine whether there was a significant difference between the number of correctly recalled name-face pairings at the cumulative test between participants in the tested not vocalise and tested vocalise conditions. This suggested that there was no significant difference between participants' performance on the cumulative test ($F(1,38) = .667$, $p = .419$, $d = .26$), hence there was no main effect of vocalisation. The Cohen's d effect size indicates that vocalisation had a small size effect on the performance of participants at the cumulative test; vocalisation therefore only slightly assisted in the recall of name-face pairings at the cumulative test.

Discussion for experiment 2

There was a significant main effect of testing in experiment 2 for both the test at list four and the cumulative test. Participants who were tested between each list of 12 name-face pairings recalled significantly more name-face pairings than participants who restudied the stimuli from lists 1-3. The first and second hypotheses were therefore accepted.

There was no significant main effect of vocalisation at list four or at the cumulative test. Participants who vocalise names at the presentation phase did not perform significantly better than participants who read the names silently. The third and fourth hypotheses were therefore rejected.

The results from this experiment show that while testing improves the performance of participant’s ability to recall name-face pairings at a list four test and on a final cumulative test, vocalisation does not significantly improve recall. It could therefore be argued that individuals do not need to vocalise names in order for them to be able to remember them better, instead they only need to test themselves throughout the presentation of name-face pairings. Alternatively restudy is another powerful strategy by itself, especially when compared to the mnemonic technique of vocalisation.

The results from experiment 2 are, as expected, similar to the results from experiment 1 in regards to the tested not vocalise and restudy conditions which suggests consistency in the performance of participants. The results from experiment 2 show a significant difference between the number of correctly recalled name-face pairings between tested and restudy participants at the cumulative test, however, once the alpha levels account for multiple comparisons in experiment 1 there is no significant difference. This inconsistency is most likely due to a much larger standard deviation in the tested condition in Experiment 1. Therefore further research should be conducted in order to establish what is causing such a high degree of variance in the tested condition compared to the other conditions. The effect sizes in experiment 2 are similar to that in experiment 1, the results from experiment 2 again suggests that testing facilitates the learning and recall of information.

The effect sizes for correct recall and intrusion rates were much lower when investigating the effect of vocalisation compared to the effect sizes when testing was investigated. This suggests that while testing results in higher recall of name-face pairings, vocalisation does not. Although participants who vocalised names performed slightly better at the list four and cumulative test this was not significant and the effect size was small. Additionally the standard deviations were large which suggests that correct recall overlapped between the tested not vocalise and tested vocalise conditions, future research is required to further assess the effect of vocalisation on the ability to correctly recall name-face pairings.

As was identified in the first experiment, the testing effect has again been shown to be an extremely valuable learning tool in cases where individuals are required to learn many name-face pairings. Participants who were tested between the presentations of stimuli performed better on a recall test on a final list and also on a cumulative test than participants who were required to restudy lists. Testing should therefore be done to ensure that individuals are able to best retain information in the context of name-face pairings.

The literature which investigates the effect of vocalisation on performance of recall has largely been dominated by word lists being used as stimuli. Many studies which have found that vocalising improves recall have used word lists as stimuli with no other additional to be remembered information (such as faces), such as the studies conducted by MacLeod et al (2010), Rackie et al (2014)) and Gathercole and Conway (1988). The present experiment however is a paired associate task where participants were required to learn two pieces of information in combination (participants are required to remember a name which is matched to a face) which is more complex than remembering one piece of information such as items in a word list. Kellogg (2012) stated that vocalising names allows for more elaborate encoding of them, it means that there is an additional phonemic code of the processing of each name. This could strengthen the memory trace of a name and also give an additional node for recall. Name-face learning tasks are largely ignored in the vocalisation literature. One study which did investigate name-face recall and vocalisation was Pariante (1990) which provided findings which were consistent with the present experiment, that vocalisation of names does not significantly improve the recall of name-face pairings when tested.

Pariatne (1990) is the only study to the researcher’s knowledge which has investigated name-face pairings where names are presented in different modes such as auditorily presented names, vocalising names, reading names silently and mouthing names. Pariante (1990) determined that participants who silently read names performed better than other participants. Pariante's (1990) study involved participants learning only eight name-face pairs, this is dissimilar to the present experiment because 48 name-face pairings had to be learned in the present experiment. It may be interesting to learn whether the findings from Pariante's (1990) study in relation to recall and the mode of presentation of names can be applied to a more substantial memory task, where participants are required to learn multiple lists of name-face pairings. It could be argued that this may explore more ecologically valid methods of name presentation. In real life individuals are likely to hear someone say either their own name or another person's name in comparison to reading a written name, the person learning the name may then repeat the name back to the individual.

The production effect outlined by Ozubko, Gopie and MacLeod (2011) is an important theory in the area of vocalisation and memory. It states that vocalised items may be remembered more because they are distinctive when compared to silently read items. The present study included participants to either read all of the names from every list silently (tested not vocalise) or required them to read all of the names from every list aloud (tested vocalise). Thus, participants were not given an opportunity to distinguish between read silent and read aloud names. Future research into the area of name-face recall and vocalisation could incorporate a further variable where participants are required to learn half of the names aloud and half of the names silently to determine whether this assists with the learning of information. This would mean that the read aloud names were distinctive against read silent names and if this generated similar results to experiments which used word lists as stimuli (e.g. Ozubko et al, 2011), participants may be able to recall more read aloud than read silent names as they would have specific memory of speaking each name aloud and hearing it.

Although there is much research in the word list stimuli domain which suggests the vocalising items improves later recall, this was not founded in the present experiment. There have however been some conflicting findings which suggest that vocalising material does not assist an individual in remembering it, such as the findings from Murray, Leung and McVie's (1974) study. Some research has suggested that it is not the act of vocalisation which makes it an effective cue to distinctiveness (Quinlan & Taylor, 2013). Quinlan and Taylor (2013) stated that vocalising words in a normal tone does not aid an individual's memory for information. They stated that increasing retention of information in memory is dependent on other distinct elements such as volume and pitch. They identified a larger production effect for those participants who sang items when compared to participants who said the items loudly and who said the items in a normal tone. Future research could therefore investigate whether the performance of recall for name-face pairings increases depending on the pitch and volume of names spoken. This could, however lack ecological validity as in a real life situation it is very unlikely someone would sing another person's name.

General Discussion

The current study aimed to identify simpler strategies to improve memory for name-face learning as complex strategies identified in the literature are not appropriate to implement whilst conversing with others. The demands of such mnemonics are not appropriate in the real world where there are competing pressures for attention. Similar to the findings of Patton (1994), Morris, Fritz, Jackson, Nichol and Roberts (2005) demonstrated that mnemonic techniques encouraging semantic elaborations of a name (e.g., trying to remember the surname "Cook" by linking it to the profession of being a cook) failed to demonstrate any benefit when implemented whilst the participant was engaged in conversation. Although the improvement in memory was replicated under laboratory conditions, the authors demonstrated that performance was actually slightly depressed compared to a control condition not given any instructions on how to learn names when they attempted to implement the technique in real world settings.

Similarly to many laboratory experiments, the present research lacks ecological validity however it could be argued that the present study is more ecologically valid than previous research which has investigated the effect of testing on recall using word lists stimuli such as Nunes and Weinstein (2012) and Carpenter, Pashler, Wixted and Vul (2008). It is also more ecologically valid than some research conducted on vocalisation which has used words as stimuli such as Ozubko et al (2011) and Rackie et al (2014). It is unlikely that an individual would be required to remember a series of unrelated words in everyday life, therefore the present experiment and similar research such as Weinstein et al (2011) which used name-face stimuli is more ecologically valid because the majority of people will have to learn others' names. The findings of this research has many applications to people who need to know a large number of faces such as those who work with many people i.e. teachers or people who work in security or the police force.

The stimuli presented to participants in the present study were shown on a computer screen and the images of the faces were also static, if an individual was required to learn someone's name in real life they would likely see them moving, additionally it is possible for people to look different in a photograph than they look in real life. Future research could therefore involve video footage of people to improve ecological validity. In the present study the faces which participants aimed to remember were presented individually on the computer screen to participants, however in real life situations such as work events or social situations people would see several people at the same time, it may therefore be more difficult to remember and recall individuals' names. To improve ecological validity in the field of testing and recall, future research could investigate how well people learn names of individuals they have recently met in a naturalistic setting.

In a real life scenario it would be expected for individuals to remember name-face pairings over an extended period of time, this is dissimilar to the present experiment where there were merely minutes between learning and recall. There has been research in the education literature such as studies by Carpenter et al (2008) and Wenger, Thompson and Bartling, (1980) which has investigated the notion of testing and forgetting over a period of days and weeks; however there is limited research in delayed testing using name-face stimuli. It could therefore be argued that future research on name-face recall could introduce a further variable to assess whether time after exposure to information affects the performance of recall.

To conclude, this experiment supports previous literature which has suggested that testing between lists of stimuli improves the recall of name-face pairings. Testing is a simple activity that individuals can implement in everyday situations to assist their learning of individual's names. This study aimed to incorporate a further variable, vocalisation in order to discover whether both testing and vocalisation could improve recall of name-face pairings. There was, however, no significant effect of vocalisation found when participants were tested a final presented list and at a cumulative test. Further research could be conducted in the area of names presented in different modalities to further understand the utility of this mnemonic in regard to name-face learning.

References

Agarwal, P.J., Karpicke, J.D., Kang, S.H.K., Roediger, H.L., & McDermott, K.B. (2008). Examining the testing effect with open and closed-book tests. *Applied Cognitive Psychology*, 22(7), 861-876. doi: 10.1002/acp.1391.

Anderson, J.R., & Bower, G.H. (1972). Recognition and Retrieval Processes in free Recall. *Psychological Review*, 22, 261-295. doi:10.1037/h0033773.

Arenberg, D. (1977). The Effects of Auditory Augmentation on Visual Retention for Young and Old Adults. *Journal of Gerontology*, 32(1), 192-195. doi:10.1093/geronj/32.2.192.

Arnold, K.M., & McDermott, K.B. (2013). Free recall enhances subsequent learning *Psychonomic Bulletin & Review*, 20(3), 507-513. doi:10.1080/20445911.2011.532207.

Baddeley, A.D., Gathercole, S., & Papango, C. (1998). The Phonological Loop as a Language Learning Device. *Psychological Review*, 105(1), 158-173. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/9450375.

Bennett, R.W. (1975). Proactive Interference in Short-Term Memory: Fundamental Forgetting Processes. *Journal of Verbal Learning and Verbal Behaviour,* 14,123-144. doi: 10.1016/S0022-5371(75)80060-0.

Blumenfeld, R.S., & Ranganath, C. (2007). Prefrontal Cortex and Long-Term Memory Encoding: An Integrative Review of Findings from Neuropsychology and Neuroimaging. *Neuroscientist*, 13(3), 280-291. doi: 10.1177/1073858407299290.

Brédart, S., Brennen, T., Delchambre, M., MCNeill, A., & Burton, A.M. (2005). Naming very familiar people: When retrieving names is faster than retrieving semantic biographical information. *British Journal of Psychology*, 26, 205-214. doi: 10.1348/000712605X38378.

Brown, C., & Lloyd-Jones, T.J. (2003). Verbal overshadowing of multiple face and car recognition: effects of within- versus across-category verbal descriptions. *Applied Cognitive Psychology*, 17(2), 183-201.doi: 10.1002/acp.861.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77(3), 305-327.

Calderwood, L., & Burton, A.M. (2006). Children and adults recall the names of highly familiar faces faster than semantic information. *British Journal of Psychology*, 97, 441-454. doi: 10.1348/000712605X84124.

Carpenter, S.K., & DeLosh, E.L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, 19(5), 619-636. doi: 10.1002/acp.1101.

Carpenter, S.K., & DeLosh, E.L. (2006). Impoverished Cue Support Enhances Subsequent Retention: Support for the Elaborative Retrieval Explanation of the Testing Effect. *Memory & Cognition*, 34(2), 268-174. doi:10.3758/BF03193405.

Carpenter, S.K., Pashler, H., Wixted, J.T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition*, 36(2), 438-448. doi: 10.3758/MC.36.2.438.

Carpenter, S.K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 35(6), 1563-1569. doi:10.1037/a0017021.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, 20(6), 633–642. doi:10.3758/BF03202713.

Castel, A.D., Rhodes, M.G., & Friedman, M.C. (2013). Predicting memory beneifts in the production effect: the use and misuse of self-generated distinctive cues when making judgments of learning. *Memory and Cognition*, 41(1), 28-35. doi: 10.3758/s13421-012-0249-6.

Cohen, G. (1990). Why is it difficult to put names to faces?. *British Journal of Psychology*, 81(3),287–297. doi: 10.1111/j.2044-8295.1990.tb02362.x.

Cohen, G., & Burke, D.M. (1993). Memory for proper names: A review. *Memory*, 1(4), 249-263. doi: 10.1080/09658219308258237.

Cull, W.L. (2000). Untangling the Benefits of Multiple study Opportunities and Repeated Testing for Cued Recall. *Applied Cognitive Psychology*,14(3), 215–235. doi:10.1002/(SICI)1099-0720(200005/06)14:3<215::AID-ACP640>3.0.CO;2-1.

De Haan, E. H. F., Appels, B., Aleman, A., & Postma, A. (2000). Inter- and intramodal encoding of auditory and visual presented material: Effects on memory performance. *The Psychological Record*, *50*, 577–586. Retrieved from: http://opensiuc.lib.siu.edu/tpr/vol50/iss3/10.

Divis, K.,& Benjamin, A.S. (2014). Retrieval speeds context fluctuation: Why semantic generation enhances later learning but hinders prior learning. *Psychonomic Society, Inc*, 42, 1049-1062. doi:10.3758/s13421-014-0425-y.

Forrin, N. D., MacLeod, C. M., & Ozubko, J. D. (2012). Widening the boundaries of the production effect. *Memory and Cognition*, 40(7), 1046–1055. doi:10.3758/s13421-012-0210-8

Gathercole, S.E. & Conway, M.A. (1988). Exploring Long Term Modality Effects: Vocalization best Leads to Retention. *Memory & Cognition*, 16(2), 110-119. doi:10.3758/BF03213478.

Golding, J.M., Long, D.L., & MacLeod, C.M. (1994). You can't always forget what you want: Directed forgetting of related words. *Journal of Memory and Language*, 33(4), 493-510. doi:10.1006/jmla.1994.1023.

Greenlee, J.D.W., Jackson, A.W., Chen, F., Larson, C.R., Oya, H., Kawasaki, H, Chen, H., & Howard, M.A. (2011). Human Auditory Cortical Activation during Self-Vocalization. *Plos One*, 6(3). doi:10.1371/journal.pone.0014744.

Gruneberg , M.M., & Herrmann, D.J. (1997). *Your Memory for Life!: Develop, Improve and Retain Lifetime Memory Skills*. Richmond: Blandford Press.

Haist, F., Shimamura, A.P., & Squire, L.R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18(4), 691-702. doi:10.1037//0278-7393.18.4.691.

Higbee, K. (2001). *Your Memory: How It Works and How to Improve It*. New York: Marlowe & Company.

Hale, A.D., Skinner, C.H., Williams, J., Hawkins, R, Neddenriep, C.E. (2007). Comparing comprehension following silent and aloud reading across elementrary and secondary students: Implication for curriculum-based measurement. *The Behavior Analyst Today*, 8(1), 9-21. doi: 10.1037/h0100101.

Helder, E., & Shaughnessy, J. J. (2008). Retrieval opportunities while multitasking improve name recall. *Memory*, 16, 896–909. doi:10.1080/09658210802360611.

Hourihan, K.L., & MacLeod, C.M. (2008). Directed forgetting meets the production effect: Distinctive processing is resistant to intentional forgetting. *Canadian Journal of Experimental Psychology*, 62(4), 242-246. doi: 10.1037/1196-1961.62.4.242.

Icht, M., & Mama, Y. (2015). The production effect in memory: a prominent mnemonic in children. *Journal of Child Language*, 42, 1102-1124. doi:10.1017/S0305000914000713.

Jang, Y., & Huber, D. E. (2008). Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 112-127.doi: 10.1037/0278-7393.34.1.112.

Jones, E. E., & Lockhart, A. V. (1919). A study of oral and silent reading in the elementary schools of Evanston. *School and Society*.

Juel, C., & Holmes, B. (1981). Oral and silent reading of sentences. *Reading Research Quarterly*, 16(4), 545-568.doi: 10.2307/747315.

Kampf, M., Nachson, I., & Babkoff, H. (2002). A serial test of the laterality of familiar face recognition. *Brain and Cognition*, 50(1), 35-50. doi: 10.1016/S0278-2626(02)00008-8.

Karpicke, J., Butler, A., Roediger, H. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own?. *Memory,*17(4), 471-479. doi: 10.1080/09658210802647009.

Karpicke, J.D., & Roediger, H.L. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249-255.doi: 10.1111/j.1467-9280.2006.01693.x.

Kellogg, R.T. (2012). *Fundamentals of Cognitive Psychology*. California: Sage.

Kuo, T., & Hirshman, E. (1997). The Role of Distinctive Perceptual information in Memory: Studies of the Testing Effect. *Journal of Memory & Language*, 36, 188-201. doi: 10.1006/jmla.1996.2486.

Landauer, T.K., & Bjork, R.A. (1978). Optimum rehearsal patterns and name learning. *Practical aspects of memory*, 625-632.

Lau E.F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: deconstructing the N400. *Neuroscience*, 9, 920-933. doi: 10.1038/nrn2532.

Lindsay, D.S., Allen, B.P., Chan, J.C.K., & Dahl, L.C. (2004). Eyewitness suggestibility and source similarity: Intrusions of details from one event into memory reports of another event. *Journal of Memory and Language*, 50, 96-111. doi: 10.1016/j.jml.2003.08.007.

Lustig, C., May, C.P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, 130(2), 199-207. doi: 10.1037//0096-3445.130.2.199.

MacLeod, C. M., Gopie, N., Hourihan, K.L., Neary, K.R., & Ozubko, J.D. (2010). The production effect: Delineation of a phenomenon. *Journal of experimental psychology: Learning, memory, and cognition,* 36(3), 671-685. doi: 10.1037/a0018785.

Markman, A. (2010). *Say it Loud: I'm Creating a Distinctive Memory.* Retrieved 17th September 2015 https://www.psychologytoday.com/blog/ulterior-motives/201005/say-it-loud-i-m-creating-distinctive-memory.

Matthews, T. D., Smith, R. E., Hunt, R. R., & Pivetta, C. E. (1999). Role of Distinctive Processing During Retrieval. *Psychological Reports*, 84, 904-916. doi:10.2466/pr0.1999.84.3.904.

McCarty, D.L. (1980). Investigation of a visual imagery mnemonic device for acquiring name-face associations. *Journal of experimental Psychology: Learning, Memory and Cognition*, 6(2), 145-155. doi: 10.1037/0278-7393.6.2.145.

McWeeny, K.H., Young, A.W., Hay, D.C., & Ellis, A.W. (1987). Putting names to faces. *British Journal of Psychology*, 78(2), 143-149. doi: 10.1111/j.2044-8295.1987.tb02235.x.

Morris, P.E., Jones, S., & Hampson, P. (1978). An imagery mnemonic for the learning of people's names. *British Journal of Psychology*, 69(3), 335-336.doi: 10.1111/j.2044-8295.1978.tb01663.x.

Morris, P.E., Fritz, C.O., & Buck, S. (2004). The name game: Acceptability, bonus information and group size. *Applied cognitive psychology*, 18(1), 89-104.doi: 10.1002/acp.948.

Morris, P.E., Fritz, C., Jackson, L., Nichol, E., & Roberts, E. (2005). Strategies for learning proper names: expanding retrieval practice, meaning and imagery. *Applied cognitive psychology*, 19(6), 779-798. doi: 10.1002/acp.1115.

Murray, D.J., Leung, C., & McVie, D.F. (1974). Vocalisation, Primary Memory and Secondary Memory. *British Journal of Psychology*, 65(3), 403-413. doi:10.1111/j.2044-8295.1974.tb01414.x.

Nunes, L.D., & Weinstein, Y. (2012). Testing improves true recall and protects against the build-up of proactive interference without increasing false recall. *Memory,* 20(2), 138-154. doi:10.1080/09658211.2011.648198.

Ozubko, J.D., Gopie, N., & MacLeod, C.M. (2012). Production benefits both recollection and familiarity. *Psychonomic Society,* 40, 326-338. doi:10.3758/s13421-011-0165-1

Pariate, G. (1990). *The effect of vocalized, mouthed, auditory, and silent rehearsal on elders' name-face recall*. (Master's Theses, San Jose State University). Retrieved from http://scholarworks.sjsu.edu/etd_theses.

Parr, L. (2011). The evolution of face processing in primates. *The Royal Society,* 366(1571), 1764-1777. doi:10.1098/rstb.2010.0358.

Pastötter, B., Weber, J., & Bäuml, K.H.T. (2013) Using testing to improve learning after severe traumatic brain injury. *Neuropsychology,* 27(2), 28-285.doi: 10.1037/a0031797.

Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval During Learning Facilitates subsequent Memory Encoding. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 37, 287–297. doi:10.1037/a0021801.

Patton, M. Q. (1994). Developmental evaluation. *Evaluation Practice*, 15(3), 311-319. doi: 10.1016/0886-1633(94)90026-4.

Quinlan, C.K., & Taylor, T.L. (2013). Enhancing the Production Effect in Memory. *Memory*, 21(8), 904-915. doi:10.1080/09658211.2013.766754.

Rackie, J.M., Brandt, K.R., & Eysenk, M. W. (2014). Interaction between mode of learning and subjective experience: Translation effects in long-term memory. *Memory*, 23(3). doi:10.1080/09658211.2014.88670.

Rivolta, D., Palermo, R., & Schmalzl, L. (2013). What is overt and what is covert in congenital prosopagnosia?. *Neuropsychol Rev*, 23(2), 111-116.doi: 10.1007/s11065-012-9223-0.

Roediger, H.L., & Karpicke, J.D. (2006). The power of testing memory: basic research and implications for educational practice. *Perspectives on Psychological Science*, 1, 181-210. doi:10.1111/j.1745-6916.2006.00012.x.

Roediger, H.L., & Karpicke, J.D. (2008). The critical importance of retrieval for learning. *Science*, 319(5868), 966-968. doi:10.1126/science.1152408.

Rosenthal, J., & Linnea, C.E. (2011). Pronouncing new words aloud during the silent reading of text enhances fifth graders' memory for vocabulary words and their spellings. *Springer*, 24(8), 921-950. doi: 10.1007/s11145-010-9239-x.

Schmidmaier, R., Ebersbach, R., Schiller, M., Hege, I., Holzer, M., & Fischer, M.R. (2011). Using electronic flashcards to promote learning in medical students: retesting versus restudying. *Medical Education*, 45, 1101-1110. doi: 10.1111/j.1365-2923.2011.04043.x.

Shohamy, D., & Adcock, A.R. (2010). Dopamine and adaptive memory. *Trends in cognitive science*, 14(10), 464-472. doi: 10.1016/j.tics.2010.08.002.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008).Testing during study Insulates against the build up of Proactive Interference. *Journal of Experimental*

Psychology. Learning, Memory, and Cognition, 34, 1392–1399. doi:10.1037/a00130821392.

Takashima, A., Segers, E., Fernándezb, G., Verhoevena, L. van den, B. (2013). Neural correlates of testing effects in vocabulary learning. *NeuroImage*, 78, 94-102. doi:10.1016/j.neuroimage.2013.03.071.

Tell, P.M. (1971). Influence of Vocalization on Short-Term Memory. *Journal of Verbal Learning and Verbal Behaviour*, 10, 149-156. doi:10.1016/S0022-5371(71)80006-3.

Tomlinson, T.D., Huber, D.E. Rieth, C.A., & Davelaar, E.J. (2009). An Interference Account of cue-Independent Forgetting in the no-think Paradigm. *Proceedings of the National Academy of Sciences*, 106(37), 15588-15593. doi:10.1073/pnas.0813370106.

Tong, Z., Han, C., Luo, W., Wang, X., Li, H., Luo, H., Zhou, J., Qi, J., & He, R. (2013). Accumulated hippocampal formaldehyde induces age-dependent memory decline. *PubMed*, 35(3), 583-596.doi: 10.1007/s11357-012-9388-8.

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

Weinstein, Y., McDermott, K.B., & Szpunar, K.K. (2011). Testing protects against proactive interference in face–name learning. *Psychonomic Bulletin & Review,* 18(3), 518-523. doi:10.3758/s13423-011-0085-x.

Wenger, S. K., Thompson, C.P., & Bartling, C. A. (1980). Recall facilitates subsequent recognition*. Journal of Experimental Psychology: Human Learning & Memory*, 6(2), 135-144. doi: 10.1037/0278-7393.6.2.135.

Wissman, K.T., Rawson, K.A., & Pyc, M.A. (2011). The interim test effect: Testing Prior Material can Facilitate the Learning of New Material. *Psychonomic Bulletin & Review,* 18(6), 1140-1147. doi:10.3758/s13423-011-0140-7.

Zaromb, F.M., & Roediger, H.L. (2010). The testing effect in free recall is associated with enhanced organisational processes. *Memory & Cognition*, 38(8), 995-1008. doi: doi:10.3758/MC.38.8.995.

Appendix

List 1	List 2	List 3	List 4
Aaron	Aiden	Brian	Andrew
David	Alfie	Callum	Ashley
Dylan	Arthur	Charlie	Billy
Jacob	Calvin	Daniel	Bradley
Johnny	Edward	Jamie	Darren
Kevin	Harry	Jason	Declan
Martin	Jimmy	Jason	Ethan
Matthew	Lewis	Michael	Francis
Phillip	Marcus	Nathan	Gavin
Robert	Richard	Patrick	Jerry
Simon	Stephen	Peter	Jordan
Terry	Stewart	Thomas	Roger