Enactment and Retrieval

Daniel J. Peterson

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology

Chapel Hill
2009

Approved by:

Neil Mulligan

Kelly Giovanello

Peter Ornstein

Abstract

Daniel J. Peterson: Enactment and Retrieval

(Under the direction of Neil Mulligan)

A number of memory phenomena are modulated by experimental design, with the effect (e.g. of bizarreness, generation, perceptual interference) occurring in recall for mixed-list but not pure-list designs. These effects have other similarities and have been treated in common theoretical frameworks, some focusing on encoding others on retrieval. The typical paradigm for examining design effects confounds encoding and retrieval contexts, making it difficult to compare these accounts. McDaniel et al. (2005), using a new paradigm, concluded that retrieval processes contribute to the bizarreness effect. This paradigm was applied to the related enactment effect. Participants were presented with two pure study lists, and later recalled the lists separately (inducing pure retrieval sets) or recalled the lists together in a single test (inducing a combined or mixed retrieval set). In three experiments, the combined recall condition consistently failed to enhance the size of the enactment effect. The results provide no support for the retrieval account of these two variables but are generally consistent with encoding accounts.

Acknowledgements

First and foremost, I would like to thank my advisor, Dr. Neil Mulligan, for his direction and guidance which were an invaluable part of this project.

I would also like to thank the other members of my committee, Dr. Peter Ornstein and Dr. Kelly Giovanello, for their time and helpful comments to make this project a success.

Table of Contents

List of Tables

Table

List of Figures

Figure

Chapter 1

Introduction

Beginning in the early 1980's researchers discovered that movement can enhance memory for action phrases. In these studies, participants either performed simple tasks (such as tapping a desk or breaking a pencil), watched an experimenter perform the same task, or read a description of the task (e.g. Engelkamp, 1998). Tasks carried out by the participant (participant performed tasks or SPTs) resulted in better memory for the events compared with both experimenter performed tasks (EPTs) and written instructions, a phenomenon dubbed the enactment or SPT effect.

Independently of one another, three separate labs, Engelkamp & Krumnacker (1980), Saltz & Donnenwerth-Nolan (1981), and Cohen (1981) discovered this phenomenon. Engelkamp and colleagues (Engelkamp & Krumnacker, 1980; Engelkamp & Zimmmer, 1984; Engelkamp & Zimmer, 1985) discovered the SPT effect while researching methods to improve memory. Engelkamp & Krumnacker (1980; as cited in Engelkamp & Zimmer, 1985) had participants listen to instructions to perform a task (no performance), imagine performing a task, observe an experimenter performing the task, or perform the task themselves. Actually performing the act led to the best recall, simply listening to instructions resulted in the worst recall, and imagining or observing the task led to intermediate and comparable levels of performance (Figure 1). Engelkamp & Zimmer (1985) theorized the improvement in recall was because enactment led to representation of the performed act in three ways: at the verbal level (e.g. a written instruction to break a pencil), the imagery level (seeing the pencil) and the motororic level (i.e. the tactile sensation and kinesthetic feedback of breaking the pencil). This motor program processing leads to increased performance over simply listening to an instruction or watching the EPT.

Engelkamp's (e.g. Engelkamp & Krumnacker, 1980; Engelkamp & Zimmmer, 1984; Engelkamp & Zimmer, 1985) theory was motivated, in part, by Paivio's (1971; 1995) dual code theory in which Paivio claims that conceptualization of an item at both a verbal and visual level leads to improved memory relative to encoding only one or the other. The motor component provides another type of coding in addition to the verbal and visual modalities improving memory beyond observation, Engelkamp theorized, offering the first explanation as to why enactment leads to improved recall.

Like Engelkamp, Saltz & Donnenwerth-Nolan (1981) studied enactment as a means of improving memory. In their study, participants either listened to or performed actions, with performance leading to improved recall. The authors framed their explanation for the memory improvements in terms of distinctiveness. The motor component is crucial in that it creates a more distinctive memory trace which is more accessible during recall.

Cohen (1981) used enactment to investigate the generality of 'memory laws' constructed on research with verbal materials. He wanted to determine if putative laws of memory applied to memory for actions as opposed to words. Specifically, he was interested in whether or not serial position effects or levels of processing manipulations, two well established phenomena from the verbal-learning literature, held during enactment.

Typical serial position effects can be observed when participants are presented with and later asked to recall a list of items. Recall often reveals a U-shaped curve of performance such that items presented at the beginning and end of lists are remembered better than items in the middle, the primacy and recency effects, respectively. However, the explanations behind memory improvements for earlier and later items differ. The standard account says that the items at the end of the list are recalled well because they are still in short term memory having just been presented. Items at the beginning of the list, too, are recalled well because participants have had the most time to rehearse them without later items interfering, making them more likely to be transferred to long term memory and subsequently recalled. As such, the primacy effect is taken as an indicator of intentional (or strategic) rehearsal processes. The items in the middle of the list are not recalled as well because they do not have the benefit of being in short term memory and were inadequately rehearsed, as previous items monopolized rehearsal processes governing transfer to long term memory.

Across a series of experiments, Cohen (1981) had participants listen to lists of words, perform a series of actions or watch an experimenter perform the same actions. The enactment data exhibited the standard recency effect, but interestingly across all three experiments, there was no significant primacy effect as there was in the verbal and EPT conditions. Post-experimental questioning of the participants helped give insights to the potential cause. Participants reported that during enactment, they were not trying to actively remember the actions they were performing, as opposed to the verbal tasks in which they repetitively rehearsed the items as much as they could. The failure to actively rehearse the performed actions presumably explained why there was no primacy effect. Cohen took this as evidence for enactment as a nonstrategic encoding process, a view later research generally supports (e.g. Engelkamp, 1998; Nilsson, 2000 cf. Kormi-Nouri, 1995, 2000; Helstrup, 1986).

Perhaps more interesting than a failure to find a primacy effect was Cohen's (1981) failure to find a levels-of-processing effect. Typically recall is better when participants are instructed to attend to the 'deep' features of an item, such as its meaning, than if required to attend to 'shallow' features of an item, such as its graphemic appearance (Craik & Tulving 1975). A similar manipulation had no effect on enactment, and current research substantiates the notion that enactment is insensitive to levels of processing manipulations (Helstrup, 1987; Nilsson, 2000). Cohen (1989) claimed that this provides further evidence that enactment is nonstrategic. He claimed enactment to be the most efficient means to encode a stimulus- any attempt to further improve encoding would be fruitless (Cohen, 1989).

The enactment effect: Current conceptualization

In explaining why enactment improves memory, researchers today have come to a general consensus on two factors (Nilsson, 2000). The first is the notion that enactment is a nonstrategic process, as first proposed by Cohen (1981). Support for this claim comes from Cohen's (1981) original study in which he failed to find primacy effects, plausibly because participants were not actively trying to rehearse the actions they performed. In a similar vein, he identified that rate of presentation (up to 20 seconds between each item) did not have any effect on enactment (Cohen, 1989). For verbal presentation, when participants have more time between items, they have more time to rehearse them and therefore improve recall. However, during enactment the extended interval between items had no impact. Further evidence comes from Backman, Nilsson, & Kormi-Nouri (1993) in which participants were given divided attention tasks while listening to verbal instructions or carrying out to-be-remembered actions. When attention was divided, recall of verbal instructions was impaired though memory for actions remained unaffected, as would be expected of an automatic encoding process.

Population dissociations also support the nonstrategic claim. Backman & Nilsson (1984) showed that older adults were significantly worse at remembering verbal tasks compared with younger adults but that there were no significant differences with enacted phrases[1]. Similarly, Cohen & Bean (1983) showed that though mentally retarded adults could not recall action phrases as well as normal adults, there was no difference in recall for SPTs.

The second and more critical factor is that the motor component is an independent subsystem[2] in memory and plays a crucial role in producing an SPT advantage (e.g. Engelkamp, 1998). Much of the research supporting this view comes from Engelkamp's lab in which he repeatedly demonstrates that the motor component is necessary to obtain benefits in recall (e.g. Engelkamp & Zimmer 1985; Engelkamp, 1998). In their original study, Engelkamp & Krumnacker (1980, as cited in Engelkamp & Zimmer, 1985) demonstrate that actual performance leads to the highest levels of recall compared with listening, observing and imagining.

To demonstrate that the motor component is truly independent, Engelkamp &Zimmer (1984) showed that it could in fact be isolated from verbal instruction. In this study participants were primed with a motor action that they either heard or performed and were then presented with a different target action. With respect to the motor actions, the prime and

[1] More recently, Ronlund, Nyberg, Backman and Nilsson (2003) found age differences in enactment but indicate that for adults between the ages of 35-55 the effect may be mediated by educational differences

[2] Here, the terms system or subsystem are used as Engelkamp uses the term. This is not to imply it meets the criteria of a formal independent memory system as Schacter, Wagner and Buckner (2000) outline.

target were either congruent, (e.g. paint the door and wave the handkerchief) or incongruent (e.g. paint the door and beat the carpet), and participants were to determine which. If participants had acted out the prime, they were faster to determine if the actions were congruent compared with if they had simply heard the prime. Engelkamp & Zimmer took this to mean that information represented by the motor component goes beyond that of simple verbal instructions. These results indicate that motor information is available during encoding and may guide retrieval.

Additional evidence comes from Zimmer & Engelkamp (1985) in which participants were given motor phrases to enact while viewing a potentially interfering video. The authors here make the distinction between motor actions (that which the enactment literature speaks of, e.g. breaking a pencil) and kinematic movement in which the person is stationary (observing movement from one location in space to another, e.g. the train goes by). When participants enacted phrases while viewing video clips of motor actions, fewer of these phrases were recalled than when they viewed clips of kinematic movement. In other words, while kinematic movement left enactment relatively unaffected, viewing motor actions interfered with participants' ability to encode enacted phrases. The authors took this as further evidence of the independence and importance of the motor component with respect to enactment arguing that because the motor component is an independent system it is resistant to sources of interference which are not also motoric in nature.

More direct evidence comes from Engelkamp, Zimmer, Mohr & Sellen (1994) in which enactment was manipulated not only during encoding, but at recall as well. Participants encoded a list of phrases either verbally or through enactment and were later given a recognition test. Critically, at test, participants either read the phrases or acted out each phrase before giving an old/new decision. When participants acted out a phrase at test which they had previously enacted, recognition was better than if they had only enacted the phrase during encoding, a phenomenon labeled the reenactment effect. Importantly, there was no significant benefit for enacting a phrase at test which was verbally encoded at study. These results can be understood in terms of encoding specificity (e.g. Tulving & Thomson, 1973) which states that the greater the extent to which study and test conditions match, the greater the benefits will be seen at retrieval. Here, assuming an item is enacted at study, its additional enactment at test improves recall.

In further support of the motor component theory, research has helped rule out other alternative explanations of the enactment advantage. Mulligan & Hornstein (2003) found an SPT advantage even when participants were blindfolded, indicating that visual imagery or visual feedback alone does not drive the effect. Further, Kormi-Nouri (2000) found that even blind participants could demonstrate an SPT effect. Similarly, several researchers have required participants to interact with imaginary rather than actual objects (e.g. Engelkamp & Zimmer 1984; Kormi-Nouri 2000). Still, there is a benefit for enacted phrases refuting the possibility that it is simply the tactile feedback of the object in one's hand that improves memory.

Enactment and the item-specific vs. relational account

As reviewed, enactment improves memory for phrases across a variety of contexts. Tests of recognition (e.g. Engelkamp et al., 1994), free recall (e.g. Engelkamp & Dehn 2000), and cued recall (e.g. Kormi-Nouri, 1995) all demonstrate that enacted phrases are remembered better than verbal presentation or EPTs. However, enactment does not always improve memory. Research has demonstrated that in certain contexts enactment can produce no benefit or even a negative effect on recall (e.g. Backman, Nilsson & Chalom, 1986; Steffens, 1999). These limitations are best understood in terms of the item-specific vs. relational framework. When learning a list of items, participants do not merely encode each individual item independently from one another; associations are also formed between the items. Hunt & McDaniel (1993) outlined the item-specific vs. relational framework to differentiate these two types of processing. Item-specific processing can be defined as processing features of a stimulus that are unique to that item, increasing the item's distinctiveness in memory. Relational processing can be thought of as the analyses of shared features, such as the connections made between cues and targets, and between individual items and the list as a whole.

Consider a study list consisting of a series of word pairs (e.g. hot-cold, north-south, etc.). Item-specific processing can be understood as the degree to which you attend to the particular features of each item, in this case, each word individually. Relational processing, conversely, can be further broken down into several subcomponents. The connections or associations made between the cue and target (in this case, antonyms) can be thought of as cue-target relational processing. The connections made between each word pair in the study list are considered inter-item relational processing. In the example given, perhaps you make the connection that it is hot in the south and cold in the north. Inter-item relational processing also refers to the connections one makes between each item and the list as a whole, in this case how you relate one particular antonym word pair to all of the other antonyms presented in the list.

Different types of memory tests rely differentially on these types of processing. Several tests rely heavily on relational processing. Tests of order memory, in which participants are required to reconstruct the order in which items were presented, rely heavily on inter-item relational processing. Cued recall tests (e.g. hot- ?) rely on cue-target relational as well as item-specific processing. Free recall relies both on relational information to help delimit the potential responses and item-specific information to accurately select among these potentially correct items. In contrast, recognition tests (deciding whether an item is old or new) do not require the generation of potential responses, so relational information is less important. As a result item-specific processing dominates on this type of test (Hunt & McDaniel, 1993).

Following this, stating that enactment improves recall does not paint the entire picture, as it simply lumps item-specific and relational processing together. In regards to item-specific processing, research has clearly demonstrated that enactment serves to improve it because of the consistent effects demonstrated with recognition memory (e.g. Engelkamp et al., 1994; Engelkamp and Dehn, 2000). Engelkamp and colleagues (Engelkamp, Zimmer & Mohr 1990) were the first to support this claim and researchers now generally agree with the notion that improvements in item-specific processing drive the enactment effect (e.g. Steffens, 1999; Steffens, Buchner , & Wender, 2003; Nilsson, 2000).

Research on relational processing is not as clear. There is agreement on one aspect of relational processing, the relation between the verb and noun. As reviewed earlier the associations made between the cue and target are thought to be one form of relational processing. In the example given previously, relational processing helps strengthen the connection between ‘hot’ and ‘cold’, which facilitates recall of the target ‘cold’ when the cue ‘hot’ is presented in isolation (i.e. on a cued recall test). In terms of enactment, a parallel can be drawn with regard to the verb and noun of an action phrase. For example, if a participant were to be asked to break the pencil, relational processing would help bind ‘break’ and ‘pencil’ such that in a cued recall test, if given ‘pencil’ participants would be better able to recall ‘break’ as its corresponding verb as opposed to other plausible alternatives, such as ‘write.’ Enactment serves to strengthen this association between verb and object more so than simply listening to verbal instructions (Nilsson, 2000).

There is not the same consensus, however, with other measures of relational memory. Only a few years after its initial discovery, Backman et al. (1986) reported that although enactment may improve recall of individual items, inter-item relational processing was actually impaired. Participants were given lists of actions to perform with varying degrees of relatedness between the items on each list (e.g. similar motor actions with objects, or similar shapes or textures of the objects themselves). Category clustering, defined as the degree to which similar items are grouped at recall, is dependent on inter-item relational memory providing a measure of how well a participant has grouped a related set of items on a list. Clustering scores in free recall revealed that regardless of how well items were related to one another, participants’ ability to organize the lists of actions together was worse after enacting the items relative to observation.

With regard to order memory, Steffens et al. (2003) has shown that when participants are asked to perform a related set of actions (organizing a back pack) their ability to recall the individual steps is worse than when they observe someone else perform the tasks. Engkelkamp and colleagues (e.g. Engelkamp, Zimmer & Mohr, 1990; Zimmer and Engelkamp 1989) originally maintained that though relational processing is not improved as a result of enactment, it is not hindered. Recently their views have slightly changed. Engelkamp & Dehn (2000) conducted a study investigating memory for enacted phrases in long (24 items) and short (eight items) lists. Participants were given a list of actions to perform or observe in pure lists (composed entirely of SPT or EPT phrases) or mixed lists (a combination of both). Memory was assessed through recognition memory (a measure of item-specific encoding), order memory (one types of relational encoding) and free recall (a test affect by item-specific and relational information). The results led the authors to three general conclusions. (1) Enactment enhanced item information indicating recognition is universally improved by enactment regardless of design type (e.g. mixed vs. pure lists, long or short lists). (2) Order reconstruction was better for EPTs relative to SPTs, but only in pure list presentations and only for shorter lists. In pure lists of EPTs, relational memory is enhanced leading to these improvements, though mixing the lists disrupts relational processing. When the lists are too long (i.e. longer than 20 items) relational processing (specifically order processing) became too demanding, and the reconstruction benefits seen with pure EPT lists disappeared. (3) In free recall, the SPT advantage generally was greater in mixed lists relative to pure lists (more on this later). This second point is most relevant to the order memory debate. Although the authors concede that EPTs leads to better order memory, they contend this is only true with regard to short lists; longer lists reveal no differences between the conditions.

In thinking of relational processing as a whole, Steffens (1999) proposes that the degree to which whole-list relational processing is improved depends on the congruence between the verb-object relations and whole-list relations. As previously outlined, enactment has been shown to strengthen the association between verb and object. If the same associations are stressed for all the items, whole-list relations will be improved, but if not, whole-list relations will be hindered. For example, participants may be presented with a list of phrases to be enacted: pick the apple, pick the cherry, pick the plum. In this case enactment should draw attention to the attributes that these objects share and therefore the category to which they belong, fruits. We would therefore expect whole-list relational processing to be improved, facilitating recall for both the verb and object. In contrast, consider a participant who is presented with the phrases: lift the horse and dress up the cat. Enacting the phrase ‘lift the horse’ should draw attention to the horse’s considerable weight, while enacting ‘dress up the cat’ should draw attention to the domesticity of cats. The one similarity that is shared, the fact that horses and cats are both animals, is not emphasized. Therefore, with such incongruent material, whole-list relational processing should be hindered by enactment, leading to a negative enactment effect. In both of these instances, inter-item relational processing is important, but the degree to which this information relates to the list as a whole, whole-list relational information, is critical.

Methodological concerns

An important issue of contention among researchers is the use of real vs. imaginary objects at study. Some researchers have chosen to employ a more ecologically valid scenario using actual objects (e.g. Mulligan & Hornstein, 2003) whereas others ask participants to carry out the action with imaginary objects (e.g. Kormi-Nouri, 2000). In his original work Engelkamp found no difference between real and imaginary objects (Engelkamp & Zimmer, 1985). Kormi-Nouri (2000) found no difference between using real vs. imaginary objects, though others have noted that the use of real objects leads to greater improvements in memory relative to imaginary objects (Ratner & Hill, 1991). Research now indicates that both imaginary and real objects result in enactment benefits (Nilsson, 2000) though the magnitude of improvement may not be equivalent. Studies using actual objects often yield a larger SPT effect than those employing imaginary objects, (Steffens et al., 2003).

Steffens et al. (2003) addressed another concern with regard to enactment, the effect of environmental cues. In most studies investigating action memory, phrases involving items that would later be removed at test (such as a pencil for the command, “break the pencil”) were intermixed with phrases using objects (e.g. the desk for, “tap on the desk”) or one’s own body parts (e.g. “pat your head”) which remained present at test. The authors noted that researchers never control for the fact that these latter actions have a definite (albeit a subtle) cue during retrieval (in this example the desk or your head) while other actions such as “break the pencil” do not. Though the cues would also be present for EPT or verbal phrases, their presence would differentially benefit enacted phrases because of the increased binding between the verb and object. The authors argued that participants may be using these environmental cues to remember many of these actions which could explain, at least in part, why enacted phrases are better recalled.

To address this question Steffens et al. (2003) had participants perform or observe actions, controlling for this potentially significant confound. Across conditions, they found participants tended to recall only the items for which there was a cue at retrieval. These results were particularly interesting, because if substantiated, they could imply that the enactment effect is nothing more than an artifact of poorly designed recall scenarios for which subtle cues are not controlled. However, Steffens, Buchner, Wender & Decker (2007) found support for an alternative explanation. In the previous study the presence of cues for previously enacted items was suppressing participants' ability to recall other action phrases. The authors conclude that the potential confound cannot explain the entire enactment effect, but maintain it merits attention in studies of enactment.

One additional issue that merits mention is the question as to whether or not physical movement is actually necessary to see improvements in memory. As reviewed earlier, most researchers agree that the motor component is crucial in finding an enactment effect (e.g. Engelkamp, 1998; Nilsson, 2000; Cohen, 1989; Steffens et al., 2003). Kormi-Nouri (1995, 2000), however, proposes a radically different view of action memory. He theorizes that use of objects as well as actual movements make no difference, so long as the participant imagines performing the action. According to Kormi-Nouri, motor components cannot explain these improvements in memory; rather enactment (as he calls it, though participants need not actually enact anything) increases the degree of self-involvement. This greater degree of self-involvement leads to a more effective strategy and rehearsal leading to elevated levels of recall. As such, he challenges the assumption that enactment is nonstrategic (e.g. Cohen, 1981; Nilsson, 2000).

The most compelling support for this claim comes from a study (Kormi-Nouri, 2000) in which participants perform tasks with real or imaginary objects, imagine performing tasks with and without real objects present or listen to verbal tasks (no objects present). It is important to note that this paradigm is very similar to Engelkamp & Krumnacker's (1980, as cited in Engelkamp & Zimmer, 1985) original study where they found a hierarchy of benefits in recall in which actual enactment led to the best recall, followed by imagining, then verbal instructions. Kormi-Nouri (2000) however, found no differences across encoding condition; regardless of condition there was a relatively standard improvement of enactment compared to verbal tasks. The crux of his argument stems from the demonstration of an enactment effect when participants are not actually enacting anything; they simply imagine performing a task. Clearly, if this is the case the motor component is not explanatory in enactment. Kormi-Nouri offers no explanation as to why his results differed from Engelkamp and Krumnacker's (1980) when the studies were so similar, and researchers have yet to replicate his results.

Integrating enactment with other memory phenomena

The explanations offered for the enactment effect can all be thought of as encoding accounts: specialized processing during encoding which produces later benefits at retrieval. McDaniel & Bugg (2008) outline how enactment can be thought of in relation to other memory phenomena with similar encoding explanations. Specifically, the authors delineate how the sometimes counterintuitive effects demonstrated with generation (improved memory for generated material), perceptual interference (improved memory for items presented only very briefly), bizarreness (improved memory for bizarre or uncommon sentences), enactment, and word frequency can all be understood in terms of an item-order account.

For these five manipulations, the effect of the variable on recall depends on the list composition. For example, when participants are required to generate an item themselves, they typically show increased recall for the item compared to having simply read it (e.g. Mulligan & Lozito, 2004). However, this, like all of the other mentioned manipulations, only holds true if generate items are interspersed with read items (a mixed list presentation). In contrast, if participants are presented with a pure list of items to be generated, memory typically is no better than a pure list of read items (Mulligan & Lozito, 2004).

McDaniel & Bugg (2008) propose an item-order explanation to account for these pure vs. mixed list effects. A variant of the item-specific vs. relational view (Hunt & McDaniel, 1993) the item-order account focuses on one specific type of relational information, order memory. Like the original item-specific vs. relational account, it specifies that item specific as well as relational processing aid in free recall. Relational information provides the structure to guide retrieval, while item specific information allows for the discrimination of list items from items not presented in the experimental context. The account more specifically claims that participants rely on order information, one type of relational processing, during free recall because it provides an organizational structure to support the search for target items.

In regards to the mixed vs. pure list distinction, the nature of the list composition differentially affects item-specific and order processing. Each variable (e.g generation, perceptual interference, etc.) contrasts an unusual and common condition. Unusual stimuli, such as items to be generated, attract more attention for interpretation resulting in greater item-specific processing. This enhancement in item specific processing, however, comes at the cost of order processing. This is thought to be because we have a limited amount of resources with which to encode an item. Therefore, there is necessarily a trade-off such that the enhancement of item specific processing limits the resources available for relational processing. Common stimuli, such as read items, attract less processing of item characteristics, allowing for greater processing of the inter-item associative processing thought to underlie order memory. During mixed list presentations of both generate and read items, the disruption of order processing caused by the presentation of generate items uniformly hinders recall of both read and generate items. However, the generate items uniquely benefit from the enhancement in item specific processing, leading to the recall benefits seen with generated material. In pure lists of generate items, the reduction of order processing roughly counter balances any item specific improvements resulting in recall levels comparable to that of read items.

Because recognition is mostly sensitive to item-specific and not relational processing, this view predicts normal effects in recognition for mixed and pure lists (in contrast to free recall). This is true of all the manipulations. In contrast, with direct measures of order memory (i.e. test of order reconstruction) unusual conditions should be worse off relative to the common conditions for all variables. Specifically order memory should be best with pure lists of common stimuli, followed by mixed lists and lastly by pure lists of unusual stimuli. These patterns have generally been found (e.g. Nairne, Riegler & Serra, 1991; Engelkamp & Dehn, 2000; McDaniel & Bugg, 2008) which supports the theory and underscores the commonality across the five variables.

Encoding vs. retrieval accounts

Often, distinctiveness has been invoked to explain these list effects (e.g. Hunt & Worthen, 2006). All of the relevant effects mentioned contrast a common with an unusual condition. This unusual condition is thought to make items more distinctive in memory, and thus more likely to be recalled. In generation, when read and generate items are intermixed, the generated items are relatively distinct from the more familiar read items, explaining why memory is improved for the generated material. In contrast, in a pure list of generate items, there is no relative distinctiveness, so memory does not improve.

This does not explain, however, why generated material would be considered distinct in a mixed list presentation with equal numbers of read and generate items. The answer lies in the notion of how distinctiveness is defined. Schmidt (2007) outlined two ways in which something can be considered distinct. Primary distinctiveness can be thought of as something that is distinctive given real world expectations. Consider the sentence, "he got in the car and turned on the mustard." Given the context of the sentence, one would expect a final word such as radio, so 'mustard' in this context would be defined as primarily distinct because it violates our expectations given what we know about the real world. On the other hand something can be locally distinctive given its context. Schmidt defines this as secondary distinctiveness- something that is noticeably different relative to the items around it (spatially or temporally). For example, one green dot among an array of red dots would have the properties of secondary distinctiveness.

In a mixed list presentation, there are often equal numbers of the common and unusual items. Both sets have a relatively equal amount of secondary distinctiveness, but it is the additional primary distinctiveness that makes unusual items such as generated material recalled better.

Building on this distinctiveness framework, McDaniel, Dornburg, & Guynn (2005) proposed an explanation for these mixed list vs. pure list effects radically different from the usual encoding accounts. According to their proposed retrieval account, an item is distinct only when compared with non-distinctive items in retrospect. Specifically, the authors use retrieval to explain why bizarre sentences are recalled more than normal sentences.

In a typical design, mixed list presentation results in a mixed retrieval set. Retrieval set here is defined as the set of potential items that may be recovered from memory. In a mixed list presentation participants read both common and bizarre sentences. In this case, the retrieval set would consist of all of the sentences read during the study portion, both common and bizarre. Importantly, the retrieval set reflects the information encoded at study. In this mixed list presentation the unusual material is relatively distinct explaining why an effect is found. In a pure list presentation there is also a pure retrieval set, meaning nothing is relatively distinct, which is why no effect is demonstrated. The typical design, however, cannot differentiate between an encoding account (like the item-order account) and a retrieval account because the mixed vs. pure encoding is confounded with the mixed vs. pure retrieval set. To appropriately test this, it is necessary to disentangle this confound which McDaniel et al. (2005) attempted to do.

In their study (McDaniel et al., 2005) participants were presented with two pure lists of bizarre or common sentences (order of list presentation was counterbalanced across participants) separated by a distracter task. For example, a participant may be presented with a pure list of common sentences (e.g. the maid wiped the ammonia off the table), followed by an unrelated distracter, then a pure list of bizarre sentences (e.g. the maid licked the ammonia off the table). At test participants were placed in one of two retrieval conditions. In the combined recall condition, participants were instructed to recall any of the sentences presented in either list during a five minute free recall session, whereas in the separate recall condition participants were given 2.5 minutes to recall any of the items presented in list 1, then 2.5 minutes to recall any of the items in list 2 (again, order of list recall was counterbalanced across participants).

For participants in the separate list recall there was no significant bizarreness effect, bizarre and normal sentence recall was roughly comparable. According to both the item order and retrieval account, this is to be expected. In thinking of the item-order account, the item-specific gains seen with the bizarre items are offset by the disruptions with relational processing, which is why memory is no better than that for the common sentences. According to the retrieval account, since both lists (and both retrieval sets) are homogenous, there is nothing relatively distinct, so memory for each list should be comparable.

Interestingly, when participants were instructed to recall the lists together, significantly more bizarre than common items were recalled. The authors argue that when participants were asked to recall the lists together, the resulting retrieval set included both lists. Only then did the bizarre sentences become distinctive. These results provide compelling evidence for the retrieval hypothesis because extant encoding views cannot account for this effect as in each recall condition, the encoding conditions were identical. The only factor which varied across conditions was the manner in which the items were retrieved (that is, the nature of the retrieval set) offering a plausible explanation for the bizarreness effect.

Enactment: Encoding or Retrieval Phenomenon?

McDaniel et al. (2005) provide an interesting alternative explanation for why manipulations such as bizarreness, which had traditionally been conceived as encoding phenomena, improve memory. Perhaps these manipulations do not elicit any specialized encoding; it may be instead that an item is only distinct when compared with non-distinctive items in memory. The current study aims to test the generality of McDaniel et al.'s (2005) findings with respect to enactment. As reviewed earlier (McDaniel & Bugg, 2008) bizarreness and enactment, along with other manipulations, behave similarly in response to experimental manipulations and have been similarly explained with encoding explanations. It is therefore reasonable to test whether this retrieval hypothesis, too, can be thought of as something characteristic of all these manipulations (bizarreness, enactment, generation, word frequency, and perceptual interference) or if this is a unique property of the bizarreness effect.

The present experiments will closely follow McDaniel et al. (2005). Participants will be presented with two pure lists, one of SPTs and one of control items separated by a distracter task. Participants will then either perform two separate recall tests (one list followed by the other) or recall both lists at once. According to the retrieval hypothesis proposed by McDaniel et al. (2005) there should be an interaction such that the enactment effect is larger for the combined recall relative to the separate recall. Conversely, according to encoding accounts (such as the item-order hypothesis, McDaniel & Bugg, 2008) because no factors vary at encoding, there should be no differences between the separate and combined recall.

Chapter 2

Experiment 1

Experiment 1 is a replication of McDaniel et al. (2005) with action phrases. All of the critical encoding and retrieval conditions are directly replicated. Because all of the encoding conditions are identical across participants, of critical importance is the comparison between participants in the combined recall condition and participants in the separate recall condition.

In this first experiment our goal was to maximize the potential SPT effect. To do so, verbal tasks served as the controls and all SPTs were enacted with real objects. The rationale for this was to give the retrieval account the most opportunity for success as these conditions give us the greatest chance of detecting an effect if there truly is one. It is important to note that this is only relevant if the retrieval account is successful in explaining distinctiveness. According to the extant encoding explanations this should make no difference and no differences will be found between the two recall conditions.

Method

Participants. Thirty-two undergraduates at The University of North Carolina at Chapel Hill served as participants in exchange for course credit.

Design. The experiment used a 2 (encoding condition: SPT vs. VT) X 2 (recall condition: combined vs. separate) design in which the encoding condition was manipulated within-participants and the recall condition was manipulated between participants.

Materials. Thirty-two simple action phrases were assembled (see Appendix) following the guidelines set forth by Steffens et al. (2003). All items involved real objects which were subsequently removed after each item presentation, leaving no retrieval cues available at test. Each of the items was constructed such that neither an object nor an action repeated throughout the entire experiment. Each of the phrases was randomly assigned to one of two lists, resulting in two lists of sixteen items each. Each list was presented in the SPT and VT conditions an equal number of times. Two additional items were placed at the beginning and end of each study list to serve as practice items and buffers, resulting in two study lists of twenty items.

Procedure. All participants were tested individually in a lab room with a small conference table. The experimenter and the participant sat facing each other so that each had a clear view of the other. First, the study phase began by presenting participants with the first list of phrases. Participants were informed that they would hear a series of action events, and their task was either to perform the action or simply to listen to the action. Each list was comprised either entirely of SPT or VT phrases. The order of the lists was counterbalanced across participants. In the SPT condition each trial proceeded as follows: First, the relevant object was brought out and placed on the table in view of the participant. Next, the action phrase was presented over the computer speaker. Then, the participant acted out the phrase, directly interacting with the object (e.g. actually breaking a pencil). After eight seconds the item was removed and the next trial began.

In the VT condition each trial followed the same procedure as the SPT condition though this time, the participant only listened to the phrases rather than acting them out. Participants were still presented with the object (so as not to confound visual presentation); they simply did not interact with it.

After the presentation of each list, participants carried out an unrelated distracter task for five minutes. Participants were presented with a sheet of math problems to solve in their head without making any notes or intermediate calculations on the paper. Participants were allowed to work the problems in any order they chose.

At test participants were randomly placed into one of two recall conditions. In the combined condition, participants were instructed to recall as many of the phrases as they could from both of the lists, both the VT and SPT phrases. This free recall test lasted five minutes, and participants were encouraged to use the entire time allotted. In the separate recall condition, participants were instructed to think back and recall all the items remembered from one list (the order of which was counterbalanced across participants) for 2.5 minutes. Following this, they were given a second recall test in which they were given an additional 2.5 minutes to think back and recall all the items remembered from the other list. In the instructions of each test, participants were reminded of the nature of the items for the to-be-recalled list (e.g., participants were asked to recall list 1 and were reminded that this was the list in which they acted out the phrases).

Results and Discussion

Table 1 presents performance on the recall phase. Both recall conditions were scored in two ways. The stringent recall score required that both the noun and verb of a phrase be recalled together to count as correct (e.g. break and pencil for *break the pencil*). The lenient score considered an item correctly recalled if either the noun or the verb were recalled. Both scores yielded the same pattern of results in this and later experiments; only the lenient recall scores are reported. The proportions recalled were submitted to a 2 X 2 analysis of variance (ANOVA), using encoding condition (SPT and VT) and recall test (combined and separate) as factors (the alpha-level was set to .05 for this and subsequent analyses). The analysis revealed a main effect for recall condition, $p < .05$, indicating higher recall in the combined than separate condition, but no significant main effect for encoding condition or interaction were found, $ps > .25$.

Before considering the more critical results, I discuss the unexpected main effect of recall condition. It is possible that the combined condition produced higher recall because it led to a functionally shorter retention interval for the list to be recalled second in the separate condition. That is, if participants in the combined condition complete the bulk of their recall in less than 2.5 minutes (which seems likely based on typical recall dynamics, e.g., Mulligan, 2007), then they have largely finished recalling both lists before the separate condition begins their second test. This may place the combined group at an advantage in terms of overall recall. However, earlier research using this paradigm (McDaniel et al., 2005; Mulligan & Peterson, 2008) has not found significant main effects of recall condition. Likewise, no main effect of recall condition was found in the subsequent experiments of the current project. Consequently, this effect will be treated as a non-replicable result and not discussed further.

According to McDaniel et al.'s (2005) analysis, the separate recall condition should be no different from traditional pure-list designs. Because pure-list designs typically do not yield an SPT effect, it is necessary to compare this condition with combined recall condition in order to induce a combined retrieval set. It should be here, according to McDaniel et al. (2005), that a significant encoding-by-recall interaction should emerge. Specifically, the SPT effect in the combined condition should be significantly larger than any SPT effect found in the separate condition. The current results provide no evidence for such an interaction.

Though these results cast doubt on the retrieval hypothesis, there is room for alternative explanations. One possible explanation stems from the concern that perhaps participants in the separate group were unable to discriminate between the two lists at recall. In other words, if these participants were spontaneously combining both lists at retrieval, then effectively, there would be no differences between the two recall groups. However examining the cross-list intrusions (recalling an item from list 1 during the list 2 recall or vice-versa) suggests this is likely not the case. Not one of thirty-two participants made a single cross-list intrusion indicating that list discriminability was not a problem.

Alternatively, it is possible that despite the instructions in the combined condition, participants are creating two separate retrievals sets, one for the SPT items and a second for the VT items. If this were the case, we have not satisfied the conditions that would constitute a true combined retrieval set. As such, a failure to demonstrate a significant interaction may not indicate flaws with McDaniel at al.'s (2005) hypothesis but rather with our experimental design. To further investigate, we examined list-based clustering in the combined condition. If participants were spontaneously performing two separate recalls, such a strategy would result in recall clustering by list. To assess this we used the adjusted-ratio-of-clustering (ARC) score (Roenker, Thompson, & Brown, 1971), a common measure of clustering which has a value of 0 for chance-level clustering, positive values for above-chance clustering, and a value of 1 for perfect clustering. An ARC score could not be computed for four of the participants because they overtly ordered their recalls on the basis of list membership. For the remaining twelve participants the mean ARC score was .11, not significantly different than zero ($|t|=1.11$, $p>.05$), indicating little evidence for a list-based retrieval strategy in the combined condition.

Chapter 3

Experiment 2

Experiment 1 produced no evidence that the SPT effect emerges at retrieval. The encoding-by-recall interaction was non-significant, and the ARC scores did not reveal list-based clustering in the combined recall condition that could alternatively explain the null interaction. Nevertheless, Experiment 2 was designed to follow up on the possibility that the two lists were so easily discriminated that participants in the combined condition were induced to treat their recall test as two separate recalls. To address this potential problem, Experiment 2 used two categories of objects, toys and office supplies, which were spread across the two lists. This new categorical structure was designed to provide a salient alternative to list-based recall for the combined recall condition. Participants in this condition could use a category-based retrieval strategy which would render a retrieval set encompassing both lists. Provided this manipulation works, the clustering of items as defined by their encoding condition should be minimized, providing a more appropriate test of the retrieval account. Ideally, nothing should change during the separate recall. That is, the list structure should still provide an effective retrieval organization when participants are explicitly instructed to recall by list.

Method

Participants. Thirty-two undergraduates at the University of North Carolina participated in exchange for credit in psychology courses.

*Design, Materials, and Procedure***.** Experiment 2 was identical to Experiment 1 except for the following modifications. Sixteen critical phrases were constructed using objects that fall within each of two categories (toys and office supplies). An additional four items from each category were constructed to serve as primacy and recency buffers for the two study lists. A total of seventeen items were used from Experiment 1 which fit into one of the two categories. The additional twenty-three items were newly constructed such that neither an object, nor action were repeated throughout the experiment. The phrases from each category were randomly divided so that half of each category appeared in list 1 and half in list 2. The study lists were randomly intermixed with the constraint that items from a single category repeated no more than twice. The presentation of the SPT and VT lists was identical to Experiment 1.

At test, the instructions for the separate condition did not change. Participants in the combined recall condition were informed about the presence of the categories across each list (a detail which had not been previously made explicit) and were encouraged to use these categories to help guide retrieval. In this way we hoped to induce participants to rely more on the categorical rather than the list structure of the items.

Results and Discussion

Table 2 presents performance on the recall phase. The proportions recalled were submitted to a 2 X 2 analysis of variance (ANOVA), using encoding condition and recall test as factors. Interestingly the results revealed a main effect of encoding such that SPT was greater than VT, $F(1, 31) = 10.92$, $p < .05$ (more on this in the general discussion). There was no main effect of recall test, and more importantly, no significant interaction between encoding condition and type of test, $ps > .6$.

Once again, we fail to demonstrate the predicted interaction which should be apparent if the SPT effect truly arises via retrieval processes. Though the results are consistent with the results of Experiment 1, our category manipulation may not have had the desired effect. By employing categories of items across the two lists, we expected that list-based clustering would be minimized in the combined condition and that category clustering would occur instead. The clustering analysis for Experiment 2 revealed a mean ARC score of .18 ($|t|=2.86$, $p>.05$) for list-based clustering and a mean ARC score of .16 ($|t|=2.24$, $p>.05$) for category based clustering indicating significant clustering for both. While we were successful in inducing participants to cluster their responses based on category membership, participants were no more likely to cluster based on category than on list membership ($|t|=.144$, $p>.05$).

Additionally we examined ARC scores to investigate the potential for a more sophisticated clustering strategy. The thirty-two critical items were divided up into four groups: (1) toys in the SPT condition, (2) toys in the VT condition, (3) office supplies in the SPT condition, and (4) office supplies in the VT condition. The mean ARC score for this four-way clustering measure was .14 ($|t|=4.42$, $p>.05$) again indicating significant clustering.

Chapter 4

Experiment 3

Both Experiments 1 and 2 produced no evidence that the SPT effect is larger in the combined than separate condition. However, we wanted to follow up on the problems with the clustering measures. The implementation of two different categories across each list had an effect on the amount of list-based clustering from Experiment 1 to Experiment 2, but participants were still clustering more on the basis of list membership relative to category membership. In Experiment 3 we attempted to resolve this problem by making category membership even more salient. To do so, within each list, we blocked all of the items on each list based on category such that the half of each list was office supplies and the second half was toys.

Method

Participants. Thirty-two undergraduates at the University of North Carolina participated in exchange for credit in psychology courses.

*Design, Materials, and Procedure***.** Experiment 3 was identical to Experiment 2 except for the following modifications. Each study list was rearranged such that the first eight critical items of that list (as well as the two preceding buffer items at the beginning of the list) were all action phrases involving office supplies. The following eight critical items (as well as the two following buffer items at the end of the list) were all action phrases involving toys. Importantly, only the order of these items has changed; no other changes were made from the Experiment 2 procedures

Results and Discussion

Table 3 presents performance on the recall phase. The proportions recalled were submitted to a 2 X 2 analysis of variance (ANOVA), using encoding condition and recall test as factors. The analysis failed to produce any significant main effects, and again, failed to produce significant interaction $ps > .4$.

For a third time, we fail to demonstrate the encoding X recall interaction one would predict given the results of McDaniel et al. (2005). Another important focus of this experiment, however, was the clustering during the combined recall. Experiment 3 blocked the presentation of the items by category to make the categorical nature of the items much more salient. Again, both list clustering and category clustering measures were significant ($ps < .05$). A paired samples t-test revealed that while still not significant, there was a trend towards more category clustering relative to list-based clustering ($|t|=1.62$, $p=.13$). Additional support for this conclusion comes from two of sixteen participants who spontaneously constructed two separate recall lists based on category. Because the items were not recalled in a continuous order, it is not possible to calculate the standard ARC score for these participants. However, these two participants clearly organized their recall around category information. As such an additional analysis was performed where these two participants were given a category clustering ARC score of 1 (an indication of perfect clustering). With this alteration, the difference between the two groups is now significant ($|t|=2.32$, $p<.05$). Finally, the category-based ARC scores in Experiment 3 are substantially higher than the category-based ARC scores in Experiment 2. All of these results are consistent with the notion that participants are clustering more on the basis of category in the present experiment. Rendering category information more salient did not change the central result: the combined recall condition did not induce (or increase) the SPT effect.

Chapter 5

General Discussion

Distinctiveness has traditionally been explained in terms of encoding processes (e.g. McDaniel & Bugg, 2008; Nilsson, 2000; Engelkamp, 1998), though the results of McDaniel et al. (2005) have placed these assumptions in question. According to these authors, distinctiveness is a phenomenon that operates at retrieval. Specifically, the retrieval set from which an item is recalled has a greater bearing on recall than factors operating at encoding. In this sense a distinctive item only becomes distinct in retrospect when compared with other non-distinctive items.

The present experiments aimed to test this hypothesis with respect to the enactment effect, one manner in which information can be rendered distinctive. Traditionally, the SPT effect (superior recall for enacted items) has been described as an encoding phenomenon. However, the results of McDaniel et al. (2005) raise the possibility that this effect is better understood as a retrieval phenomenon. All the critical factors from McDaniel et al.'s (2005) original study were replicated with efforts made to ensure participants were truly constructing either combined or separate retrieval sets to appropriately distinguish between the two study lists. According to the retrieval hypothesis, when participants were instructed to think back to both lists (forming a combined retrieval set), the enacted items should be relatively distinctive and thus more memorable, resulting in a larger SPT effect in the combined than separate condition. In contrast, extant encoding views posit that, because nothing was manipulated during encoding (the only factors which varied were at retrieval), any SPT effect should be identical across recall groups.

In looking across the three present experiments one thing is clear: instructing participants to recall both lists in one combined recall session does not yield a larger SPT effect than having them perform two separate recalls. As is shown below, combining data from all three experiments, to yield a more powerful analysis of the encoding-by-recall-condition interaction, also fails to demonstrate any difference in SPT effect between the two recall groups. As such, these data do not support the McDaniel et al.'s (2005) retrieval hypothesis.

Though by and large these results fit well with current encoding explanations, one quirky result was the presence of a main effect of encoding in Experiment 2. Here, items enacted were recalled significantly better than items which were only passively heard. Though the enactment effect is more often demonstrated with within-subject, mixed-list designs (e.g. Engelkamp & Zimmer, 198;, Engelkamp & Dehn 2000; Steffens, 1999; Steffens, Buchner & Wender, 2003), demonstration of an enactment effect in between-subjects and pure-list designs is not unprecedented (Koriat & Pearlman-Avnion, 2003). The more puzzling question stems from why this SPT effect was significant in Experiment 2 and not in Experiments 1 and 3. To look at this more closely, we wanted to see whether or not the SPT effect in Experiment 2 was significantly different relative to Experiments 1 and 3. These latter experiments both had numerical trends indicating greater recall in the SPT than VT condition, though the differences were non-significant. To answer this question, the data from all three experiments were submitted to a 2 X 2 X 3 analysis of variance (ANOVA), using encoding condition (SPT and VT) recall condition (combined and separate) and experiment (Experiments 1,2, and 3) as factors. Interestingly, there was a main effect of encoding such that SPT recall was greater than VT recall $F(1, 93) = 7.24$, $p < .01$. No other main effects or interactions were found to be significant, including the two-way interaction between encoding condition and experiment $F < 1$. The results from this analysis are important for two reasons. First, the lack of a significant interaction demonstrates that the results from Experiment 2 were not significantly different from Experiments 1 and 3. Second, the overall main effect of encoding is important to establish that our materials are sensitive enough to detect an SPT effect. Across all three experiments, a significant SPT effect was found only in Experiment 2. By pooling all three experiments together we can more confidently say our materials were appropriate for examining the retrieval hypothesis.

Admittedly, however, this is not the ideal manner in which to explain these effects. As such, a follow up experiment is currently underway which will address the issue more appropriately. In effect, Experiment 4 will be a combination of Experiments 2 and 3. Sixty-four participants will be randomly assigned to one of four conditions which correspond to each of the two conditions in the previous two experiments: (1) random list, combined recall (2) random list, separate recall (3) blocked list, combined recall (4) blocked list, separate recall. All of the stimuli will remain the same from Experiments 2 and 3 (though the items will be re-randomized) consisting of toys and office supplies. This should provide a more appropriate intra-experiment test as to whether or not the SPT effect really is greater when the category items are randomly intermixed versus blocked. Given that there is no theoretical reason to expect a larger SPT effect in the random condition, and given the results of the cross-experiment analysis, it is expected that the SPT effect will be a similar size in the two conditions. Additionally, with sixty-four subjects, we should have enough power to demonstrate an overall main effect of encoding (SPT greater than VT).

Because the primary conclusions from the present experiments are based upon a failure to reject the null hypothesis (i.e., the encoding-condition-by-recall-condition was always non-significant), one may be concerned with whether the results reflect a lack of power. To address this concern a formal power analysis was conducted to determine whether insufficient power was masking any effect. The pooled analysis from experiments 1-3 indicated no hint of an interaction, $F < 1$, consistent with the individual analyses. The pooled analysis has many more subjects (and thus more power) than the original study by McDaniel et al. (2005, Experiment 1) which found results consistent with the retrieval account. Given the size of the effect (f = .48) reported by McDaniel et al. (2005, Experiment 1), the presented combined analysis had power of approximately .99.

Assuming Experiment 4 turns out as expected, there would be an overwhelming amount of evidence to suggest that the retrieval hypothesis does not adequately explain the SPT effect. With the addition of Experiment 4, there would be four experiments consisting of one hundred and sixty participants in which there was not a hint of an interaction at any point. Conversely, the encoding account easily accounts for the data: because nothing systematically varied at encoding, the SPT effect between recall groups was the same.

The present results raise further questions as to why the results of McDaniel et al. (2005), who originally investigated bizarreness, are not extending to other similar memory phenomena. This study marks the third instance in which the retrieval hypothesis does provide an adequate explanation for recall benefits with distinctive material. Mulligan & Peterson (2008) looked at generation and perceptual interference with an experimental paradigm much like the current study. The results, again, indicated that the retrieval hypothesis was not able to account for when generation and perceptual interference did and did not improve recall across various experimental factors.

These four effects (bizarreness, enactment, generation, and perceptual interference) behave similarly to one another in response to many experimental manipulations causing some researchers to posit a common explanation underlying for all of the (McDaniel & Bugg, 2008). The fact that this retrieval explanation, which fits the data from McDaniel et al. (2005) so well, does not extend to other similar phenomena is concerning, and points to potential problems either with the conception of these phenomena having a similar underlying explanation, or to the results of McDaniel et al. (2005) in particular.

Mulligan & Peterson (2008) pointed to two possible explanations as to why they failed to extend the results from McDaniel at al. (2005). The first was the nature of the actual items used during study. In Mulligan & Peterson (2008) participants were presented with words in isolation, while in McDaniel et al. (2005) participants saw complete sentences. The authors speculated that perhaps more complex items produce a greater reliance on retrieval processes for bizarreness. Additionally, Mulligan & Peterson (2008) pointed to evidence that it is easier to form inter-item associations with random lists of simple items. Importantly, both of these speculative accounts hinge on the fact that Mulligan & Peterson (2008) used simple items as stimuli (words in isolation) while McDaniel et al. (2005) used more complex items (sentences). For the purposes of the present experiment, neither of these explanations hold, as the action phrases presented to participants are much like the bizarre sentences (in terms of complexity) used in McDaniel at al. (2005). As such, it is difficult to think of additional reasons why these different stimuli yield different results beyond simply stating that perhaps bizarreness is not as comparable to the other aforementioned variables as previously thought. Conversely, it might be useful to try and replicate McDaniel et al.'s (2005) original findings, as if these findings cannot be replicated, it speaks more to a single unique demonstration than a problem with conceptualizing these effects having a common underlying explanation.

Conclusions

The results of McDaniel et al. (2005) brought into question whether bizarreness, and by association, related phenomena such as enactment, can be conceptualized as an encoding phenomenon. The present experiments applied this new paradigm to enactment to investigate whether these conclusions extended to a related memory phenomenon. Participants were presented with two pure study lists, and later recalled the lists separately (inducing pure retrieval sets) or recalled the lists together in a single test (inducing a combined or mixed retrieval set). In three experiments, the combined recall condition consistently failed to enhance the size of the enactment effect. The results provide no support for the retrieval account of these two variables but are generally consistent with encoding accounts.

Table 1:

Experiment 1: Mean Proportion (SD) of Items Recalled as a Function of Encoding and Recall Conditions

Recall Condition	Encoding Condition	
	SPT	VT
Separate	.43 (.09)	.38 (.12)
Combined	.49 (.17)	.47 (.10)

Note: SPT = participant performed tasks, VT = verbal tasks

Table 2

Experiment 2: Mean Proportion (SD) of Items Recalled as a Function of Encoding and Recall Conditions

	Encoding Condition	
Recall Condition	SPT	VT
Separate	.54 (.13)	.47 (.11)
Combined	.58 (.13)	.49 (.12)

Note: SPT = participant performed tasks, VT = verbal tasks

Table 3

Experiment 3: Mean Proportion (SD) of Items Recalled as a Function of Encoding and Recall Conditions

Recall Condition	Encoding Condition SPT	Encoding Condition VT
Separate	.45 (.16)	.42 (.13)
Combined	.48 (.13)	.45 (.16)

Note: SPT = participant performed tasks, VT = verbal tasks

Figure 1:

Synopsis of mean recall scores of action phrases under different encoding conditions. From Engelkamp & Zimmer (1985)

Motor Programs and Semantic Memory

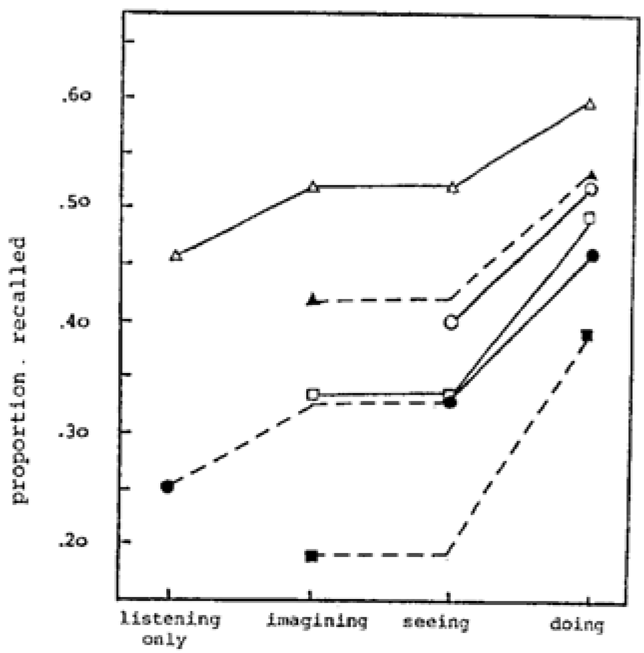

Figure 3

Synopsis of mean recall scores of action phrases under different encoding conditions.
Data from the following experiments are included:
△ Engelkamp & Krumnacker, 1980, Experiment 1;
▲ Zimmer 1984, Experiment 2;
○ Engelkamp & Zimmer, 1983, real object;
□ Zimmer & Engelkamp, 1984, Experiment 1;
● Engelkamp & Zimmer, 1983, imaginal object;
■ Engelkamp & Krumnacker, 1980, Experiment 2.
Broken lines are intrapolations.

Appendix:

Actions used in Experiment 1

ring the bell
put Easter egg together
fold the shirt
place tube standing up
click the computer mouse
shuffle the cards
rattle the keys
dial a telephone number
crumble the plastic bag
undo safety pin
stretch the elastic band
turn over the coin
pick up the battery
tear the paper
remove tape from case
roll the marble
bend the paper clip
shoot the toy gun
tie a knot using the string
close the purse

drop the pencil
push the toy car
break the match
squeeze the dog toy
wave a handkerchief
bounce the ball
pat the dog
open the book
press the stapler
unzip the zipper
seal the envelope
take cap off pen
unfold the napkin
stack the checkers
make twist tie into a V
smell the flower
wring out a sponge
shake the bottle
spin the top
turn on the flashlight

Actions used in Experiments 2 & 3

twist the elastic band
drop the pencil
Peel off label
grasp the marker
slide the binder clip
take the cap off the pen
fold the index card
turn over the clipboard
bend the paper clip
tear the paper
undo the puzzle pieces
shuffle the cards
toss the jack in the air
shake the magic 8-ball

wind the yo-yo
wave the pinwheel
bounce the ball
flip the Frisbee
pat the dog
jump the checkers
press the stapler
squeeze the hole punch
tap the ruler against the table
open the book
pull a post-it note from its stack
remove a piece of tape from its case
click the computer mouse
close the notebook

seal the envelope
pick up the eraser
shoot the toy gun
stretch the slinky
stack the blocks
rub the rubber ducky

put together the legos
place the action figure standing up
push the toy car
lift the doll
roll the marble
spin the top

References

Bäckman, L., & Nilsson, L. (1984). Aging effects in free recall: An exception to the rule. *Human Learning: Journal of Practical Research & Applications*, *3*(1), 53-69.

Bäckman, L., Nilsson, L., & Chalom, D. (1986). New evidence on the nature of the encoding of action events. *Memory & Cognition*, *14*(4), 339-346.

Bäckman, L., Nilsson, L., & Nouri, R. K. (1993). Attentional demands and recall of verbal and color information in action events. *Scandinavian Journal of Psychology*, *34*(3), 246-254

Cohen, R. L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology*, *22*(4), 267-281.

Cohen, R. L. (1983). The effect of encoding variables on the free recall of words and action events. *Memory & Cognition*, *11*(6), 575-582

Cohen, R. L. (1989). Memory for action events: The power of enactment. *Educational Psychology Review*, *1*(1), 57-80.

Cohen, R. L., & Bean, G. (1983). Memory in educable mentally retarded adults: Deficit in participant or experimenter? *Intelligence*, *7*(3), 287-298.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268-294.

Engelkamp, J., Zimmer, H. D., Mohr, G., & Sellen, O. (1994). Memory of self-performed tasks: Self-performing during recognition. *Memory & Cognition*, *22*(1), 34-39.

Engelkamp, J. (1998). *Memory for actions.*. Hove, England: Psychology Press/Taylor & Francis (UK).

Engelkamp, J., & Dehn, D. M. (2000). Item and order information in participant-performed tasks and experimenter-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 671-682.

Engelkamp, J., & Krumnacker, H. (1980). Image- and motor-processes in the retention of verbal materials. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *27*(4), 511-533.

Engelkamp, J., & Zimmer, H. D. (1984). Motor programme information as a separable memory unit. *Psychological Research*, *46*(3), 283-299.

Engelkamp, J., & Zimmer, H. D. (1985). Motor programs and their relation to semantic memory. *German Journal of Psychology*, *9*(3), 239-254.

Engelkamp, J., Zimmer, H. D., & Mohr, G. (1990). Differential memory effects of concrete nouns and action verbs. *Zeitschrift für Psychologie Zeitschrift für angewandte Psychologie*, *198*(2), 189-216.

Helstrup, T. (1986). Separate memory laws for recall of performed acts? *Scandinavian Journal of Psychology*, *27*(1), 1-29.

Helstrup, T. (1987). One, two, or three memories? A problem-solving approach to memory for performed acts. *Acta Psychologica*, *66*(1), 37-68.

Hornstein, S. L., & Mulligan, N. W. (2004). Memory for actions: Enactment and source memory. *Psychonomic Bulletin & Review*, *11*(2), 367-372.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, *32*(4), 421-445.

Hunt, R. R., & Worthen, J. B. (2006). *Distinctiveness and memory.*. New York, NY, US: Oxford University Press.

Kormi-Nouri, R. (1995). The nature of memory for action events: An episodic integration view. *European Journal of Cognitive Psychology*, *7*(4), 337-363.

Kormi-Nouri, R. (2000). The role of movement and object in action memory: A comparative study between blind, blindfolded and sighted participants. *Scandinavian Journal of Psychology*, *41*(1), 71-75.

Kormi-Nouri, R., Nyberg, L., & Nilsson, L. (1994). The effect of retrieval enactment on recall of participant-performed tasks and verbal tasks. *Memory & Cognition*, *22*(6), 723-728.

McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin & Review*, *15*(2), 237-255.

McDaniel, M. A., Dornburg, C. C., & Guynn, M. J. (2005). Disentangling encoding versus retrieval explanations of the bizarreness effect: Implications for distinctiveness. *Memory & Cognition*, *33*(2), 270-279.

Mulligan, N. W. (2001). Generation and hypermnesia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 436-450.

Mulligan, N. W., & Hornstein, S. L. (2003). Memory for actions: Self-performed tasks and the reenactment effect. *Memory & Cognition*, *31*(3), 412-421.

Mulligan, N. W., & Lozito, J. P. (2004). *Self-generation and memory.*. San Diego, CA, US: Elsevier Academic Press.

Mulligan, N. W., & Peterson, D. (2008). Attention and implicit memory in the category-verification and lexical decision tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 662-679.

Nairne, J. S., Riegler, G. L., & Serra, M. (1991). Dissociative effects of generation on item and order retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 702-709.

Nilsson, L. (2000). *Remembering actions and words.*. New York, NY, US: Oxford University Press.

Paivio, A. (1971). *Imagery and verbal processes.*. Oxford, England: Holt, Rinehart & Winston.

Paivio, A. (1995). *Imagery and memory.*. Cambridge, MA, US: The MIT Press.

Ratner, H. H., & Hill, L. (1991). The development of children's action memory: When do actions speak louder than words? *Psychological Research/Psychologische Forschung, 53*(3), 195-202.

Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L. (2003). Recall of Participant-Performed Tasks, Verbal Tasks, and Cognitive Activities Across the Adult Life Span: Parallel Age-Related Deficits. *Aging, Neuropsychology, and Cognition*, *10*(3), 182-201.

Saltz, E., & Donnenwerth-Nolan, S. (1981). Does motoric imagery facilitate memory for sentences? A selective interference test. *Journal of Verbal Learning & Verbal Behavior, 20*(3), 322-332.

Schacter, D. L., Wagner, A. D., & Buckner, R. L. (2000). *Memory systems of 1999.*. New York, NY, US: Oxford University Press.

Schmidt, S. R. (2007). Unscrambling the Effects of Emotion and Distinctiveness on Memory. In *The foundations of remembering: Essays in honor of Henry L. Roediger, III.* (pp. 141-158). New York, NY, US: Psychology Press.

Steffens, M. C. (1999). The role of relational processing in memory for actions: A negative enactment effect in free recall. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *52*(4), 877-903.

Steffens, M. C., Buchner, A., & Wender, K. F. (2003). Quite ordinary retrieval cues may determine free recall of actions. *Journal of Memory and Language*, *48*(2), 399-415.

Steffens, M. C., Buchner, A., Wender, K. F., & Decker, C. (2007). Limits on the role of retrieval cues in memory for actions: Enactment effects in the absence of object cues in the environment. *Memory & Cognition*, *35*(8), 1841-1853.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352-373.

Zimmer, H. D., & Engelkamp, J. (1985). An attempt to distinguish between kinematic and motor memory components. *Acta Psychologica*, *58*(1), 81-106.