WHAT IS IN AN INSTANCE? PRACTICE CONTEXT EFFECTS

A dissertation submitted
to Kent State University in partial
fulfillment of the requirements for the
degree of doctor of philosophy

By

Nicolas Jon Wilkins

August, 2013
Dissertation written by
Nicolas J. Wilkins
B.A., The University of Montana, 2004
M.A., Kent State University, 2010
Ph.D., Kent State University, 2013

Approved by

Katherine A Rawson, Chair, Doctoral Dissertation Committee
Jeffrey A. Ciesla, Members, Doctoral Dissertation Committee
William E. Merriman,
Christopher Was,
Brian Huot,

Accepted by
Maria S. Zaragoza, Chair, Department of Psychology
James L. Blank, Dean, College of Arts and Sciences
TABLE OF CONTENTS

LIST OF FIGURES....................................................................................................................... v
LIST OF TABLES.......................................................................................................................... vii
ACKNOWLEDGEMENTS............................................................................................................. viii
SUMMARY................................................................................................................................... 1
INTRODUCTION.......................................................................................................................... 2
  Automaticity Research, a Brief History.................................................................................. 4
CURRENT RESEARCH............................................................................................................... 11
  What is Encoded During Practice?......................................................................................... 12
  Hybrid Theory......................................................................................................................... 20
INTRODUCTION EXPERIMENT 1.................................................................................................. 22
METHOD EXPERIMENT 1............................................................................................................ 26
  Participants and Design........................................................................................................... 26
  Materials................................................................................................................................. 26
  Procedure................................................................................................................................. 27
  Results................................................................................................................................... 28
INTRODUCTION EXPERIMENT 2.................................................................................................. 38
METHOD EXPERIMENT 2............................................................................................................ 42
  Participants and Design........................................................................................................... 42
  Materials and Procedure........................................................................................................ 42
  Results................................................................................................................................... 43
INTRODUCTION EXPERIMENT 3.................................................................................................. 51
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>METHOD EXPERIMENT 3</td>
<td>53</td>
</tr>
<tr>
<td>Participants and Design</td>
<td>53</td>
</tr>
<tr>
<td>Materials and Procedure</td>
<td>53</td>
</tr>
<tr>
<td>Results</td>
<td>55</td>
</tr>
<tr>
<td>GENERAL DISCUSSION</td>
<td>60</td>
</tr>
<tr>
<td>Implications for Memory-Based Processing Theories of Automaticity</td>
<td>61</td>
</tr>
<tr>
<td>FUTURE DIRECTIONS</td>
<td>68</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>70</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>75</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1 Experiment 1 mean response times as a function of set size group, blocks of trials, and Practice and Transfer Phase. Response times are further separated into novel item and target item blocks in the Transfer Phase. Error bars equal standard error.

2 Experiment 1 percentage of trials correct as a function of set size group and blocks of trials during transfer. Percentages of trials correct are further separated into novel item and target item blocks. Error bars equal standard error.

3 Experiment 1 addend slopes of trials correct for target items as a function of set size group and mini blocks of trials during transfer. Error bars equal standard error.

4 Experiment 2 mean response times as a function of set size group, blocks of trials, and experimental phase. Response times are further separated into novel item and target item blocks in the transfer session. Error bars equal standard error.

5 Experiment 2 percentage of trials correct as a function of set size group and blocks of trials during transfer. Percentages of trials correct are further separated into novel item and target item blocks. Error bars equal standard error.

6 Experiment 2 addend slopes of trials correct for target items as a function of set size group and mini blocks of trials during transfer. Error bars equal standard error.
Experiment 3 mean response times as a function of set size group and blocks of trials. Error bars equal standard error.

Experiment 3 response times for speeded recognition task as a function of Hit, Miss, Correct Rejection, and False Alarm. Response time are further separated by set size group. Values are mean response times in milliseconds. Error bars equal standard error.
LIST OF TABLES

1 Experiment 3 mean proportion of Hits, False Alarms, and Corrected Recognition as a function of set size group.

A1 Materials for Experiments 1 and 2.

A2 Example of materials for Experiment 3 speeded recognition test.
ACKNOWLEDGMENTS

This project would not have been possible without the assistance, guidance, and support of several individuals. Specifically, I thank Katherine Rawson for advising me on this project. I would also like to thank the members of my committee: William Merriman, Jeffrey Ciesla, Christopher Was, and Brian Huot. Finally, I would like to thank Brynn Rachel Wilkins and Harper Joy Wilkins for their love and support.
SUMMARY

Memory-based processing theories of automaticity assume that shifts from rule-based processing to retrieval-based processing underlie practice gains of speed of responding. However, these theories are silent on the influence of practice context. The current work proposed a hybrid attention/memory-based processing theory in which reduction of what stimulus features are attended to during practice influences what information is encoded and thus available to be retrieved. To test the hybrid theory, the current work examined the influence of set size on initial practice gains and on performance during a transfer task. In three experiments, participants performed alphabet arithmetic verification in an initial practice phase. Results at the end of practice were consistent with the hybrid theory’s assumption that learning to reduce attention of stimulus features was more likely for participants who practiced a small set of items versus a large set of items. In Experiments 1 and 2, practice was immediately followed by a transfer phase in which trials of novel items were intermixed with practiced target items. In Experiment 3, practice was immediately followed by a speeded recognition task. Notably, in Experiments 1 and 2’s transfer phases and Experiments 3’s recognition task, results for target item accuracy were inconsistent with the original hybrid theory. However, target item accuracy results were also inconsistent with previous research on the set size effect. Thus, although current memory-based processing theories require modification to account for the current results, the hybrid theory remains unverified.
INTRODUCTION

One of the most intuitive and robust findings in cognitive psychology is that when practicing a set of cognitive problems, responses become quicker and more accurate. Examples of this phenomenon in the cognitive psychology literature are many and cover a wide range of tasks (e.g., Anderson, Fincham, & Douglass, 1999; Logan, 1988; McAndrews & Moscovitch, 1990; Rawson, 2004; Rawson & Middleton, 2009; Rawson & Touron, 2009; Schneider & Shiffrin, 1977; Touron & Hertzog, 2004; Touron, Swaim, & Hertzog, 2007; Wilkins & Rawson, 2010). Almost equal to the range of tasks investigated in the cognitive skill acquisition literature are the number of theories accounting for speed and accuracy gains during cognitive skill acquisition (e.g., ACT-R, Anderson & Lebiere, 1998; CMPL, Rickard, 1997; EBRW, Palmeri, 1997; Information Reduction Theory, Haider & Frensch, 1996; Instance Theory of Automaticity, Logan, 1988). However, few studies have investigated the influence of practice context on gains made during practice and the durability of gains after practice. Therefore, few theories of cognitive skill acquisition account for practice context effects. Thus, empirically, the current research investigates how practice context influences the gains made during practice and the durability of initial gains after practice. Given that becoming faster and more accurate during practice of a cognitive skill would have little utility if gains are not maintained after practice, investigating factors that increase durability of practice gains is important. Theoretically, a new theory of cognitive skill acquisition is proposed and
tested. If supported, the new theory will aid in predicting how practice context influences the quality of gains made during practice and the durability of gains after practice, which may be useful in developing new cognitive-skill training methods in a wide variety of areas (e.g., reading, math, job skills, etc.). However, before further discussing the current empirical and theoretical goals, a general overview of relevant theories, findings, and key issues in cognitive skill acquisition literature are discussed.

Key issues in the area of cognitive skill acquisition literature involve accounting for practice gains of speed and accuracy. In this regard, two unique ways responses to a cognitive problem are generated have been identified. And, both methods of generating a response may become faster and more accurate with practice. Response Type 1 involves the use of sequential rules that are common to all cognitive problems of a given type (rule-based processing). Thus, practicing multiple cognitive problems of a given type increases the efficiency of rule-based processing and thus increases speed and accuracy of responses. In addition, rule-based processing efficiency may also increase during practice by combining rule-steps (e.g., see Anderson & Lebiere, 1998). Response Type 2 involves generating responses by directly retrieving the answer for a specific problem from memory (retrieval-based processing). For most cognitive skills, generating responses is faster for retrieval-based versus rule-based processing because retrieval-based processing involves only one processing step. Given that retrieval-based processing involves memory for specific items, it is not initially available during initial practice of a cognitive skill. However, the likelihood of retrieval-based processing occurring increases with repeated practice of a single item (e.g., see Logan, 1988).
In addition to previously mentioned practice gains, practice gains of speed also occur when learners shift from generating responses via the rule-based processing to retrieval-based processing. This shift from rule- to retrieval-based processing is the basis of memory-based processing (MBP) theories of automaticity. MBP theories initially emerged from being a sub-area of attention theories (e.g., Kahneman, 1973). MBP theories differ on how they account for the shift from rule-based to memory-based processing, and these differences continue to be a major debate in cognitive skill acquisition literature. Given that processing shift continues to be a large focus in the memory-based automaticity literature, a brief history of how this shift has been accounted for is presented next.

Automaticity Research, a Brief History.

Early research of automatic processing was conducted in the context of the modal model of memory (e.g., Atkinson & Shiffrin, 1968) and single-capacity models of attention (e.g., Kahneman, 1973). Given this context, early studies of automatic processing focused on identifying the properties of automaticity and how the properties of automaticity differed from the properties of non-automatic processing. Although early researchers differed on the necessary properties of automaticity, some key properties of automaticity identified were that automaticity is fast, effortless, unavailable to conscious awareness, and uncontrollable/obligatory. In contrast, the properties of non-automatic processes are the opposite of automatic processes. Non-automatic processing is slow, effortful, available to consciousness, and controllable. According to single-capacity models of attention (Posner and Snyder, 1975), automatic processes are thought to not
require attention and thus are unaffected by attentional limitations brought about by tasks that use most or all attentional resources.

Although the property-list approach was important to differentiating between controlled and automatic processes, many researchers questioned what properties were necessary and sufficient for automatic processing (e.g., see Table 1, Rawson, 2004). For example, Cheng (1985) questioned the idea that automatic processing does not require attentional resources. Indeed, many studies found evidence for attentional limitations in tasks where processes used matched the properties for automaticity (e.g., Kahneman & Chajzyck, 1983; Regan, 1981). On a separate front, the idea of single-capacity attention theories was also questioned by multiple-resource theories (e.g., Wickens, 1984), which differ from single-capacity attention theories by assuming that learners have multiple attentional resource pools. Generally, each resource pool corresponds to a different mode of processing (e.g., auditory, visual, touch, etc.). Thus, the possibility exists that automatic processes are not interfered with by non-automatic processes because they draw from different attentional resources.

Another serious challenge for property-list accounts of automaticity is that their definition for automaticity is circularly defined. Specifically, property-list theories lack a mechanism for automaticity. By defining automaticity as a list of properties, automatic processing is only identifiable post experiment, after a response was made. For example, if a completed task was responded to quickly, with little to no mental effort, required little to no attention, and responses were outside the control of participants, the processing was automatic. However, if the completed task was responded to slowly, with
mental exertion, required attention, and responses were under participant control, processing was non-automatic. In sum, property-list theories of automaticity are unable to predict the transfer from consciously controlled processing to unconscious automatic processing because they do not contain a mechanism for automatic processing.

Largely in response to these criticisms, current theories of automaticity are process-based theories that do contain a mechanism for automaticity. One key finding that moved automaticity research toward process-based accounts comes from Schneider and Shiffrin (1977). Here, automatic processing was found to only develop after repeated practice for problems in which a stimulus was consistently paired with a response. In their experiments, participants repeatedly practiced identifying the presence of target items within a trial. Target items were present in half of the trials and absent in the other half. One trial consisted of viewing multiple frames that had letters and numbers on them. For example, the first frame contained the target items (e.g., 4, 7, 8, 1) the participant would search for. Once participants felt that they had memorized these target items, they pressed a start button to continue. Upon pressing the start button, a sequence of 20 more frames was rapidly presented. The participant searched these 20 frames for targets. If one of the target items (e.g., 7) was present, they would respond “present” as soon as possible. In addition, each of the 20 frames contained four items (e.g., M, J, Q, S) and the frames were divided into three sections. Section 1 was three dummy frames that never contained target items. Section 2 was 15 frames that contained the target if it was present. Section 3 was two more dummy frames that never contained a target item. Importantly, target items were consistently mapped for half of participants and variably mapped for
the other half. If consistently mapped, the target items would never be used as a distracter symbol on other trials. For example, if 3, 5, 9, and 6 were the target items on one trial, they were never distracters on a different trial. If variably mapped, target items on one trial were used as distracters on another trial.

Initially during practice, responses were slow, controllable, and effortful (non-automatic) for all participants. However, by the end of practice, responses were significantly faster for the consistently mapped group versus variably mapped group, which suggested that shifting to automatic processing only occurred for consistently mapped targets. In order to explain why responding shifts to automatic processing for consistently mapped stimuli but not for variably mapped stimuli, Memory-based processing theories of automaticity (Logan, 1988; Palmeri, 1997; Rickard, 1997) were developed.

According to memory-based processing (MBP) theories of automaticity, speed gains during practice are due to a shift from an item-general rule-based process to an item-specific retrieval-based process. According to MBP accounts, automatic processing is retrieval-based processing. For example, when first solving the problem $24 \times 7$, one would have to compute the answer (168) using algorithmic rules of multiplication, which are item-general rules because they can be applied to any multiplication problem. After multiple exposures to the same problem, eventually one will be able to retrieve the answer directly from memory, which is item-specific because retrieving the answer to this problem will not help answer other multiplication problems (e.g., $32 \times 8$).
Although all MBP theories were designed to account for the shift from rule-based to memory-based processing, each particular theory makes different assumptions about the mechanism of transfer. Indeed, the mechanism of transfer is the key debate in contemporary research. For example, instance theory (Logan, 1988) and exemplar-based random walk theory (EBRW, Palmeri, 1997) are instance-based theories in which each encounter and response to a stimulus encodes a unique memory trace for the event. Component Power Laws theory (CMPL, Rickard, 1997) and Atomic Components of Thought-Rational (ACT-R, Anderson & Lebiere, 1998) are strength-based theories in which the initial encounter and response with a stimulus encodes a memory trace and each subsequent encounter with the stimulus strengthens the one memory trace. In addition to the ongoing debate concerning memory representation, current debate also surrounds whether learners have control over which route they use to generate responses. On one side, instance theory and EBRW assume no control over the processing route used to generate a response. On the other side, CMPL and other researchers assume that learners have at least some control over which process is used to respond (e.g., Touron & Hertzog, 2004; Wilkins & Rawson, 2011). Although these debates are a pressing matter in the contemporary cognitive skill acquisition literature, they matter less for the current research. Thus, this paper will emphasize instance theory (Logan, 1988) more than other theories because instance theory is the progenitor of memory-based processing theories and is the basis for the new theory presented and tested here.

Instance theory of automaticity (Logan, 1988) has four major assumptions. First, each encounter and response to a stimulus encodes a unique memory trace for the event.
Second, instance theory is a race model, which assumes that when a stimulus is presented the rule-based process races in parallel with the retrieval-based process. The process that wins the race is the process that generates the response. Third, instance theory assumes that every time a stimulus is attended to, a unique trace of the event is obligatorily encoded into memory. That is, encoding traces into memory is an unavoidable consequence of attention. For example, if 24 X 7 has been attended to 20 times, after the 20th encounter, 20 instances of the stimulus and response are stored in memory. Fourth, instance theory assumes that every time a stimulus is attended to, retrieval of corresponding instances from memory is obligatorily initiated.

Together, these four assumptions account for the shift from rule-based to retrieval-based processing during repeated practice of items. For example, on the first encounter with 24 X 7, the rule-based process will win the race because no traces for 24 X 7 = 168 exist in memory for retrieval. After multiple encounters with 24 X 7 however, traces are available in memory to race against rule-based processing. Each trace stored for 24 X 7 = 168 has the same mean retrieval time. However, retrieval time for each trace is stochastic; retrieval time for a given trace varies around the mean. Thus, as the number of traces stored in memory increases with practice, the likelihood that a trace for 24 X 7 = 168 will be retrieved quickly increases, in turn increasing the likelihood that retrieval will win the race with the rule-based process.

Although instance theory describes a basic mechanism to account for shift from rule-based to retrieval-based processing during practice, instance theory is silent concerning moderating factors that may influence the rate of shift in process use during
practice. In general, one goal of the current research is to extend instance theory to account for moderating factors that influence the rate of process shift.
CURRENT RESEARCH

As a reminder, the current research has two goals. First, most research on memory-based automaticity has been focused on the shift from rule-based to retrieval-based processing. Only a small number of studies have investigated what information is encoded into memory during practice of a cognitive skill (e.g., Lassaline & Logan, 1993; Logan & Etherton, 1994) and how the quality of information encoded may influence gains made during practice. Thus, the empirical goal of the current study is to extend the memory-based automaticity literature in regards to what is encoded during practice, by addressing two specific questions: First, does practice context influence what information is encoded? If so, does differential encoding during practice have consequences on later performance? Second, given the literature’s focus on the shift from rule-based to retrieval-based processing and its minimal investigation of what information is encoded during practice, current memory-based automaticity theories are unable to account for differential encoding during practice. Thus, the theoretical goal of the current research is to propose an extension of instance theory (Logan, 1988) that can account for differential encoding during practice and the consequences differential encoding may have for later performance.
What is Encoded During Practice?

Extensive investigation has been conducted on the shift to retrieval-based processing. However, few studies have investigated what information is encoded during practice of a cognitive skill. One such study (Lassaline & Logan, 1993) used a visual numerosity judgment task and found that not all features of a stimulus are necessarily encoded. After repeatedly practicing a set of visual patterns to achieve memory-based processing, performance was not affected by changing the elements composing the patterns (e.g., changing squares to circles) or by changing the color of the patterns. Rather, performance was affected by changing the shape of the patterns. Although results were consistent with the idea that not all stimulus information is encoded during practice, a mechanism accounting for what information is encoded was not provided.

Logan and Etherton (1994) established attention as a key component to what information is encoded. Their proposed explanation was derived from instance theory’s (Logan, 1988) assumption of obligatory encoding, which states “what is attended to is encoded; what is attended to is represented in memory” (Logan & Etherton, 1994, pp. 1022-1023). Specifically, according to their attention hypothesis, an instance is the encoded representation of stimulus features and the co-occurrence of stimulus features (co-occurrence refers to the consistent pairing of features within a stimulus processing episode). Attention to stimulus features and co-occurrence of stimulus features is required for encoding to occur.

Regarding the previously described visual numerosity task (Lassaline & Logan, 1993), according to this Attention Hypothesis, participants were unaffected by elements
of each pattern changing from practice to transfer because the type of element within numerosity patterns was not attended to. Rather, participants attended and thus encoded only the number of elements and overall pattern (co-occurrence of the spatial relation between elements) of numerosity patterns during practice.

In order to test the Attention Hypothesis (Logan & Etherton, 1994), participants performed a word-pair category identification task in which participants repeatedly searched two-word displays for members of a target category. For example, participants were given a target category (e.g., furniture) at the beginning of practice. After presentation of the category, practice trials of two-word display started (e.g., CHAIR-GRAPE). Participants identified if one of the words was a member of the target category as quickly and accurately as possible. Half of the items were target present and half were target absent. The set of items was repeatedly presented during practice. According to the Attention Hypothesis, if both words were attended to, an encoded instance would consist of representations for CHAIR and GRAPE, as well as a representation that CHAIR and GRAPE occurred together. Testing the attention hypothesis consisted of manipulating whether both words of a target present item were always the same on each repetition. If both words of a target present item were the same on each repetition, participants would attend to and encode co-occurrence information. And, participants would take advantage of their knowledge of co-occurrences when responding. Furthermore, if co-occurrences were encoded during practice, changes to the word pairings of target present items during transfer would cause changes in performance because participants would no longer be able to use their stored co-occurrence information to respond.
Experiment 1 investigated the assumption that learners will attend and thus encode co-occurrences. Half of the participants were in the consistently mapped (CM) group (e.g., if the target word of a word pair was “CHAIR” and the other word of the word pair was “GRAPE” for one presentation during practice, the words co-occurred on all subsequent presentations during practice). The other half were in the variably mapped (VM) group (e.g., if the target word of a word pair was “CHAIR” and the other word of the word pair was “GRAPE”, the target word “CHAIR” would never again co-occur with the word “GRAPE” during practice). According to the attention hypothesis, performance would be faster at the end of practice for the CM group than for the VM group because the difference between groups reflects the CM group’s “knowledge of consistent co-occurrence” (Logan & Etherton, 1994, p. 1024). Consistent with the attention hypothesis, response times at the end of practice were faster for the CM group than for the VM group, which suggests that the CM group encoded co-occurrences and that co-occurrence information aided retrieval.

Immediately after practice, target words were paired with a novel word during a transfer phase for all participants. Here, response times were slower for the CM group than for the VM group, which further suggests that the CM group encoded co-occurrences during practice and used this information to aid responding. When co-occurrence information was no longer useful the CM group’s performance suffered. Although Logan and Etherton offered no mechanism for why co-occurrences would cause a performance decrement in the target identification task when word pairs were altered, instance theory (Logan, 1988) assumes that performance would be worse for the
CM group than for the VM group because the CM group’s instances encoded during practice did not match stimuli during transfer. According to instance theory, only instances that are identical to the stimulus presented enter the retrieval process.

Experiment 1 established that learners attend to co-occurrences if they are present. After establishing the encoding of co-occurrences, Logan and Etherton explicitly manipulated participants’ attention towards learning co-occurrences. For example, their Experiment 3 was identical to Experiment 1 except one word of each word pair was colored green. Participants were told that target words would always be green and thus they would not need to attend to the non-colored word. Thus, according to the attention hypothesis, the CM group would not encode co-occurrences because co-occurrences would not be attended to during practice. Rather, only green words would be encoded because green words would be the only words attended to. Results supported the attention hypothesis. At the end of practice, response times were not significantly different between CM and VM groups. During transfer, response times were not significantly different between CM and VM groups. Together, the results suggest that attention influences what stimulus information was encoded.

These initial studies suggest that not all information of the stimulus is necessarily encoded and that attention is required for attention to occur. Furthermore, Logan and Etherton’s (1994) study manipulated attention with one factor, explicit instructions/color of target items. Minimal research on other factors that may influence what is encoded during practice has been conducted. Thus, the empirical goal for the current research is to examine another factor that may influence what is encoded during practice. The current
research investigates if practice context influences attention towards a stimulus thereby affecting what is encoded during practice. In addition, if practice context does affect what is encoded during practice by influencing attention, a further empirical goal is to investigate consequences of differential encoding during practice for later performance.

Complementing the empirical goal for the current research, the theoretical goal of the current research is to propose an extension of Logan’s instance theory that can account for differential encoding during practice and the consequences that differential encoding may have for later performance. Although the results from Lassaline and Logan (1993) and Logan and Etherton (1994) support the attention hypothesis (what is attended to is encoded) as a viable extension of instance theory, the attention hypothesis only states what is attended is encoded. The attention hypothesis does not provide a mechanism that explains why some stimulus features are attended to and thus encoded while other features are not. Thus, the only way to know if a feature or co-occurrence of a stimulus was encoded during practice is to observe a change in performance during some sort of transfer session. Regarding the visual-numerosity verification task, the attention hypothesis does not explain why the type of element (e.g., circles or squares) within a visual pattern was not encoded. Rather, we only know that the type of elements of each pattern was not encoded because changing the elements of a pattern (e.g., from circles during practice to squares) during transfer did not result in performance changes.

Regarding the word-pair category identification task (Logan & Etherton, 1994), using English word-pairs in which readers will implicitly read from left to right and from top to bottom adds constraint to how stimuli of these tasks will be attended to and
encoded. For example, for the CM group in Experiment 1, half of the target words were presented as the first word of the word pair and half were presented second during practice and transfer. Thus, for the target words, co-occurrence of words may not have been encoded because word position was held constant throughout practice and transfer. If CHAIR was the target word, always presented with the word GRAPE, and CHAIR was the second word in the word pair on one trial, CHAIR was always the second word on subsequent trials. Thus, response times at the end of practice being faster for the CM group than for the VM group could simply reflect the CM group only attending, encoding, and responding to the first word (GRAPE) of the word pair. The VM group could not do this for trials in which the target word was the second word in the word pair because the filler word was always changing. Response times during transfer were slower for the CM group versus the VM group because response times were only for correct trials. Thus, for the CM group to respond correctly, they would need to scan the entire stimulus, and they had less practice than the VM group doing this. So, response time differences during transfer may have been due to less practice attending to the entire stimulus for the CM group versus the VM group. Further evidence that implicit rules of English reading was the cause of attention differences and thus performance differences comes from the fact that in Experiment 3 participants were instructed to only attend to colored words and to ignore the other word. Instructions were required to stop attending in the characteristic way of English readers. In sum, the attention hypothesis states that attention influences encoding. However the attention hypothesis does not answer the question of why some information is attended to while other information is not.
Although the attention hypothesis (Logan & Etherton, 1994) that comes from the obligatory encoding assumption of instance theory states that whatever stimulus information is attended to during practice gets encoded, the current attention hypothesis does not include a mechanism that specifies how attention can be influenced to cause differential encoding during practice. Thus, the theoretical goal of the current research is to extend instance theory so that it can account for factors that influence attention and in turn cause differential encoding.

As previously mentioned, instance theory (Logan, 1988) assumes a four part mechanism for explanation of transfer from rule-based to memory-based processing during repeated practice of a set of items of a given type of cognitive task. However, instance theory does not specify a mechanism for how attention is influenced and thus what information is encoded in an instance during practice. One promising account for how attention can be influenced during practice is Haider and Frensch’s (1996) information reduction theory. According to information reduction theory, over the course of practice people learn to separate task-relevant information from task-redundant (or irrelevant) information. Once the separation of relevant and redundant information is accomplished, people learn to ignore the redundant information and solely focus on the relevant information, which increases their efficiency. To be clear, information reduction theory as originally formulated is an item-general account that applies to all stimuli of the task. Also note that this account can provide post-hoc explanations for findings in the previously mentioned studies. The type of element in visual-numerosity patterns (Lassaline & Logan, 1993) was not encoded because it was task-irrelevant information.
For the CM group (Logan & Etherton, 1994), a target item presented as the second word of a word-pair was not attended and thus not encoded because it was task redundant information. The first presented filler word was all that was required to correctly perform the task and thus was possibly the only word attended and encoded in the CM group. In addition to providing an account for the previous discussed research on a visual numerosity and word-pair task, information reduction theory also provides an account for practice gains in an alphabet string verification task (Haider & Frensch, 1996).

In Experiment 2 (Haider & Frensch, 1996), participants practiced verifying the correctness of alphabet strings such as A [4] F G H. Alphabet strings were correct if the string skipped the correct number of letters in the alphabet and continued the alphabet in the correct order. During practice, the first letter to the right of the digit determined if the alphabet string was correct or not, all other letters to the right of the digit were redundant because they always continued the alphabet in the correct order. However, participants were not told about redundant information. Furthermore, the amount of redundant information was manipulated by having alphabet strings with tails of one to four letters (e.g., A [4] F to A [4] F G H I). Thus, if participants learned to ignore redundant information during practice, differences in response times between alphabet strings of different lengths should decrease. Response time differences between alphabet string lengths did decrease significantly, suggesting that participants learned to ignore the redundant letters.

In addition, in their Experiment 2, practice was immediately followed by a test block. In the test block, the tailing portions of the correct alphabet strings from practice
were made incorrect by changing a letter in the redundant part of the letter string. Thus, if participants learned to ignore the redundant information in practice, the participants would not attend to the changed letter during test and would make more errors in responding. Indeed, error rates for items that changed redundant information were significantly greater in the test block than in the last block of practice, suggesting that participants ignored the redundant information in test.

Overall, Haider and French’s results suggest that people learn to attend to task relevant information and ignore task redundant information during practice. However, information reduction theory also has limitations. Specifically, information reduction theory is concerned only with accounting for item-general, rule-based efficiency gains and has no memory-based mechanism accounting for encoding or subsequent retrieval of stimulus information.

Hybrid Theory

To briefly revisit, instance theory assumes that stimulus information attended to is encoded. However, instance theory does not specify a mechanism for what information gets attended. Thus, instance theory is currently silent in regards to how differential encoding during practice may occur. Unlike instance theory, information reduction theory provides a mechanism for how attention to stimulus information may change during practice, but is silent concerning encoding and subsequent retrieval of information. Thus, a hybrid theory combining information reduction theory and instance theory would provide a mechanism for how attention to a stimulus changes during practice and would predict that attention changes during practice would cause differential encoding.
In general, if a practice context manipulation makes redundant information easier to detect in one group than another, reduction of attended information during practice would likely be greater for one group than the other. Assuming the condition is met, according to the proposed hybrid theory, encoded information would be different between the groups. The theoretical goal of the current research is to test the predictions made by the new hybrid theory under a condition in which differential learning of redundant information exists and compare these predictions to predictions of instance theory.
INTRODUCTION EXPERIMENT 1

In Experiment 1, participants learned an alphabet arithmetic (AA) verification task. In the AA verification task, participants verify whether AA items (e.g., $A + 2 = C$) are true or false as quickly and accurately as possible. Participants solve AA problems using one of two possible processes. Learners may solve AA items by use of rules, starting in the alphabet at the first term of the item ($A$), then counting up the alphabet by the addend ($2; A, B, C$), finally comparing their answer to the third term of the item to decide whether the equation is true or false. Alternatively, after repeated practice with the same AA item, learners may retrieve the answer directly from memory.

To examine if practice context influences attending to stimuli and thus what is encoded, Experiment 1 manipulated practice set size. During practice, half of the participants were repeatedly presented with a small set of AA items and half were repeatedly presented with a large set of AA items. Importantly, all items presented in the practice phase consisted of a unique first term (e.g., the $A$ in the item $A + 2 = C$) and a unique third term. Uniqueness of terms was important because it made individual terms in each item redundant, which provided the necessary conditions for information reduction to occur. For example, if a participant was presented with $A + 2 = C$, no other AA item contained the letter $A$ in the first position or the letter $C$ in the third position during practice. Thus, although the participant at first attended to the whole item and encoded the whole item and the answer (e.g. $A + 2 = C$, TRUE), after continued practice
participants may learn to reduce attended information to just the first term and encode the partial stimulus and the response (e.g., A, TRUE).

During transfer, all participants were repeatedly presented with six practiced AA items and six novel AA items. Importantly, each novel AA item shared their first two terms with one of the practiced AA items (e.g., practiced, A + 2 = C; novel, A + 2 = D).

Although uniqueness of terms of each practice item provided the necessary condition for information reduction to occur, the set size manipulation is most important because it possibly sets up conditions in which the rate of learning to ignore redundant information is different between groups. Specifically, I assumed that learning what information is redundant will be easier and thus occur more quickly when practicing a small number of items (small set group) than a large number of items (large set group). Support for this assumption comes from explicit memory research on the list length effect. The list length effect is that recognition accuracy decreases as list length increases (see Gronlund & Elam, 1994; Strong, 1912). Although why the list length effect occurs is the topic of ongoing debate (see Kinnell & Dennis, 2011), the general idea in the list length literature is that short lists afford greater item-specific processing than long lists whereas long lists afford more contextual processing than short lists.

Given that the assumptions of the hybrid theory, if my assumption of the set size effect is correct, the hybrid theory predicts differences between the small set versus the large set group. Specifically, if rate of learning redundant information during practice is faster for the small set versus large set group, reduction of attention to just the first term of a stimulus and thus partial encoding of the stimulus will occur faster for the small set
versus the large set group. Thus, the following predictions for Experiment 1 are made on my assumption that the number of partially encoded instances during practice is higher for the small set group than for the large set group.

During the first block of practice, response times for practiced items will not be significantly different between small and large set groups because neither group will have learned to ignore redundant information. However, as practice progresses, Prediction 1 is that response times at the end of practice will be significantly faster for the small set versus the large set group. Prediction 1 reflects the idea that the likelihood of partial attention and thus partial encoding of stimuli is higher for the small set group than for the large set group. In contrast, if instance theory without modification is correct and only the number of instances in memory drives the speed and likelihood of retrieval, response times will not be significantly different between small and large set groups.

Prediction 2 concerns performance in the transfer phase and the extent to which differential encoding has consequences for later performance. If the number of partially encoded target stimuli is higher for the small set group versus the large set group, accuracy for novel items during Transfer Block 1 will be significantly worse for the small set group versus the large set group. Prediction 2 reflects the idea that incorrectly retrieving partial target instances for novel items will be higher for the small set group versus the large set group.

Prediction 3 is that response times for target items during transfer will be slower for the small set versus the large set group. Prediction 3 reflects the idea that the likelihood of reverting back to using the rules to generate a correct response will be
higher for the small set versus the large set group. For both Predictions 2 and 3, instance theory would predict no difference in accuracy or response time between the groups because the only factor driving retrieval speed and likelihood is the number of instances stored in memory for a particular item.

In Experiment 1, a third of AA items were of addend size 2 (A + 2 = C), a third addend size 3 (B + 3 = E), and a third of addend size 4 (C + 4 = G). Increasing the addend size of AA items increases the number of algorithmic steps of counting up the alphabet. Thus, if participants’ responses are computed using rule-based processing, response times will increase monotonically with addend size. However, if participants’ responses are based on retrieval from memory, response times will be unaffected by addend size because all retrieval is assumed to be a one step process. Given that each block of practice contains an equal number of items of each addend size, addend slope can be calculated by fitting a linear slope across response times for correct items of addend size 2, 3, and 4. Addend slope can be used to diagnose the type of process used to solve items. During practice, addend slope is not predicted to significantly differ between the groups and is expected to approach zero by the end of practice in both groups. Key here, according to the hybrid theory, Prediction 4 is that addend slope at the beginning of transfer will be greater for the small set group than for the large set group. Prediction 4 reflects the idea that retrieval errors due to partially encoded instances will cause the small set group to ignore their retrieved responses and revert to using the rule-based process (see, Wilkins & Rawson, 2011). Instance theory would predict that addend slope would not be significantly different between groups at the beginning of transfer.
METHOD EXPERIMENT 1

Participants and Design.

Sixty-two undergraduates from Kent State University participated for research credit in an introductory psychology class. Participants were randomly assigned to one of two practice groups (small set or large set), defined by the number of items practiced.

Materials.

Stimuli consisted of 24 alphabet arithmetic items (see Table A1 in Appendix). Half of the items were true (e.g., B + 2 = D), and half were false (e.g., G + 2 = J). False items were derived from true items, but the answer was one count further down the alphabet from the true answer. Furthermore, a third of the items were of addend size 2 (e.g., A + 2 = C), a third were of addend size 3 (e.g., B + 3 = E), and a third were of addend size 4 (e.g., R + 4 = V). The alphabet arithmetic items were divided into four sets, including a target set of six items, one filler set of 12 items, and one novel set of six items. Each set consisted of an equal number of true and false items, with an equal number of each addend size. Each item of the target set and filler set were unique, such that no item shared a first or third term with any other item in these sets. Novel items were matched with target items. If a given target item was true (e.g., F + 2 = H), its matching novel item was false (e.g., F + 2 = I), and vice versa.
Procedure.

At the beginning of the practice session, participants read instructions and then practiced three true and three false sample items for two blocks to become familiar with the task. The six examples were used only as warm-up items and were not presented to participants during practice. Upon completion of the warm-up trials, participants were given feedback concerning their average response speed and accuracy. At this time, the experimenter checked participants’ performance to ensure they had completed the warm-up items appropriately and to answer any questions about the procedure. If so, participants started the main part of the experiment.

All participants completed 30 blocks of practice trials, with order of items randomized anew in each practice block. However, to decrease the likelihood of participants having partially activated information in working memory for an item when they encountered it the next time, randomization was constrained to ensure that at least two other items occurred between trials for a given item. In the small set group, a practice block contained only the target set of items. However, in the large set group, a practice block contained the target set and the filler set of alphabet arithmetic items (the filler set was randomly mixed with the target items during each block of practice). After practice, all participants immediately completed ten blocks of transfer trials, which were randomized anew in each block. In each transfer block, all participants were presented with the practiced target items and with the novel items. Filler items were not presented during transfer.
On each trial of the practice and transfer phases, an orientation stimulus (*** ) was presented for 500 ms, followed by an AA item and two response buttons (‘‘TRUE’’ and ‘‘FALSE’’) appearing below the AA item. After participants clicked on a response button, the stimulus and the response buttons disappeared. If the response was incorrect, a red ‘‘ERROR’’ message was presented for 1000 ms, followed by a button labeled ‘‘next’’. If the participant’s response was correct, the ‘‘next’’ button was immediately presented. The participant then clicked on the ‘‘next’’ button to present the orientation stimulus for the next trial, followed by the next AA item, and so on. Response time was recorded as the time between stimulus onset and clicking one of the two response buttons. Feedback concerning average speed and accuracy was presented after every 48 trials. Participants were also given the opportunity to take a small break during the presentation of feedback. After completing the main task of the experiment, participants were asked if English was their primary language and rated their state of mental fatigue on an eleven point Likert-type scale (0 = no mental fatigue to 10 = extreme mental fatigue).

Results.

Data for two participants that performed below 75% accuracy were dropped from analyses. Except where otherwise noted, analyses of response times were conducted on correct response trials only (excluding response times less than 50ms and greater than 9000ms). To minimize the effects of outliers, raw response times were log transformed, averaged over trials for each participant, and the individual averages were then anti-log transformed (Rickard, 1997, 2007; Wilkins & Rawson, 2010, 2011).
Individual differences. Regarding fatigue, ratings of fatigue were significantly greater for the large set group (M = 6.58, SE = 0.55) than for the small set group (M = 4.50, SE = 0.54), $t(50) = 2.69$, $p = .010$. However, conclusions made from inferential tests for a priori predictions were the same for tests that controlled for fatigue versus tests that did not control for fatigue. Thus, for simplicity, reported results are from tests without fatigue reports as a control.

Regarding primary language, eight participants reported that English was a secondary language. Results of mixed factor ANOVAs for target response times during practice, target response times during test, addend slope for target items during test, and overall accuracy during test were not significantly different between primary and secondary English speakers, all $F$s < 2.2, all $p$s > .111. Thus, results reported next include both primary and secondary English speakers.

Practice session. For each participant, response times for the six target items were averaged for each block of practice and for each block of transfer. Similarly, for transfer blocks containing novel items, response times for the six novel items were averaged.

Mean response time data for target and novel items are displayed in Figure 1.
To revisit Prediction 1, if the rate of focusing on relevant information and ignoring redundant information significantly faster for the small set versus large set group, attending to part of each item and encoding partial instances would be significantly higher for the small set versus large set group. Thus, response times would be significantly faster for the small set group than for the large set group at the end of practice. Indeed, target item response times at the end of practice were faster for the small set group than for the large set group during practice (see Fig. 1). A 2 (group: small set versus large set) X 30 (practice block) mixed-factor ANOVA for target item response times resulted in significant main effects of group and block, and a significant group X block interaction, $F(1,58) = 10.32$, MSE = 12024121, $p = .002$, $F(29,1682) = 160.43$, 

Figure 1. Experiment 1 mean response times as a function of set size group, blocks of trials, and Practice and Transfer Phase. Response times are further separated into novel item and target item blocks in the Transfer Phase. Error bars equal standard error.
MSE = 316333, $p < .001$, and $F(29,1682) = 2.76$, MSE = 316333, $p < .001$ respectively. Furthermore, an independent samples $t$-test for target item response times in the last block of practice confirmed that response times were significantly faster for the small set group than for the large set group at the end of practice, $t(58) = 2.08$, SE = 115, $p = .021$, which is consistent with the hybrid theory. Importantly, the results are inconsistent with Logan’s (1988) instance theory, which would predict no significant response time differences during practice between small versus large set groups.

Transfer. Percentages of target and novel items correct for each participant were averaged for each block of transfer. Mean percentage of target and novel trials correct for each block of transfer and each practice group are displayed in Figure 2.

![Figure 2](image_url)

**Figure 2.** Experiment 1 percentage of trials correct as a function of set size group and blocks of trials during transfer. Percentages of trials correct are further separated into novel item and target item blocks. Error bars equal standard error.
To revisit Prediction 2, if the number of partially encoded instances is higher for the small versus large set group, accuracy for novel items in Transfer Block 1 will be worse for the small set group than for the large set group. Prediction 2 reflects the idea that the likelihood of incorrectly retrieving partial target instances for novel items is higher for the small set group versus the large set group. However, accuracy for novel items during the first block of transfer was higher for the small set group than for the large set group (see Figure 2). An independent samples $t$-test for percentages of novel items correct in Transfer Block 1 resulted in significantly higher accuracy for the small set group ($M = 83.9\%$, $SE = 3.5$) than the large set group ($M = 56.3\%$, $SE = 4.9$), $t(58) = 4.65$, $SE = 5.92$, $p < .001$, which is unexpected and inconsistent with the hybrid theory. However, the result is also inconsistent with instance theory. More discussion of this unexpected result will occur after the rest of Experiment 1 results are presented.

Revisiting Prediction 3, target item response times at the beginning of transfer will be significantly slower for the small set versus the large set group. A 2 (group: small set versus large set) X 10 (transfer block) mixed factor ANOVA for target item response times resulted in a significant main effect of block and a block X group interaction that approached significance, $F(9,522) = 19.35$, $MSE = 131639$, $p < .001$ and $F(9,522) = 1.89$, $MSE = 131639$, $p = .051$ respectively. Although the main effect of group was not significant [$F(1,58) = .963$], the prediction of the hybrid theory was most concerned with target item response times at the beginning of transfer. Thus, an independent samples $t$-test on target item response times during the first transfer block resulted in significantly slower response times for the small set group ($M = 1835$, $SE = 113$) than for the large set
group (M = 1488, SE = 119), t(58) = 2.12, p = .019. This pattern is consistent with the hybrid theory and inconsistent with instance theory.

![Addend slopes graph](image)

Figure 3. Experiment 1 addend slopes of trials correct for target items as a function of set size group and mini blocks of trials during transfer. Error bars equal standard error.

For Prediction 4, addend slope for practiced items would be higher for the small set versus the large set group at the beginning of transfer. Prediction 4 reflects the idea that incorrect retrieval causes reversion back to using the rule and that the likelihood of reversion to the rule is higher for small set versus the large set group. Given the small number of trials for target items of each addend size in each block of transfer, for each group, two consecutive blocks of transfer were averaged together to create mini blocks. Addend slopes computed for each mini block are reported in Figure 3. As can be seen in Figure 3, addend slope for target items during the first mini block of transfer was higher for the small set versus the large set group. A 2 (group: small set versus large set) X 5
(mini block) mixed factor ANOVA for addend slope resulted in the main effect of group approaching significance, $F(1,58) = 2.82$, MSE = 316369, $p = .10$. Although the main effect of group approached significance, the prediction of the hybrid theory was most concerned with target item addend slopes at the beginning of transfer. Thus, an independent samples $t$-test on target item addend slopes during the first mini block of transfer was conducted and resulted in no significant target item addend slope differences between the small set group ($M = 221.4\text{ms/addend}, SE = 89.5$) than for the large set group ($M = 44.7\text{ms/addend}, SE = 64.7$), $t(58) = 1.58, p = .119$. Although results did not reach the level of significance (possibly due to small number of items of each addend size and/or the large variability inherent in addend slope data), the pattern is consistent with the hybrid theory and inconsistent with instance theory.

In sum, results from Experiment 1 largely supported the hybrid theory’s predictions of the set size effect. Recall that if my assumption of the set size effect was correct, learning to ignore redundant information and thus attend to only task relevant information would occur faster for the small set group than for the large set group. Given that only attended information is encoded, the number of partial instances encoded was assumed to be greater for the small set group versus the large set group. In turn, a greater number of partially encoded target instances for the small set versus large set group would create worse performance during transfer for the small set versus large set group. Response times for target items during practice were significantly faster for the small set versus large set group. Response times for target items during transfer were significantly slower for the small set versus large set group. Although not significant, the trend was
that addend slope for target items during transfer were higher for the small versus large set group. However, accuracy for novel items during transfer was unexpectedly significantly higher for the small versus large set group. Importantly though, instance theory would predict no differences between groups on all measures during practice and transfer.

Regarding the unexpected result of novel item accuracy, one possible explanation still assumes that learning to reduce information during practice was faster for the small set versus the large set group. However, Experiment 1 predictions focused only on target items without regard of the possible influence of filler items during practice. Specifically, learning to focus attention on only task relevant information and ignore task redundant information could occur for both the small and large set groups during practice. This is because filler items were also constructed with unique first and third terms. Thus, although during practice more target item trials were partially encoded for the small set group than for the large set group, the total number of trials partially encoded during practice may have been fewer for the small set group than for the large set group. For example, the small set group completed 180 trials (6 items X 30 blocks). If they started to partially encode during Block 6 of practice and all further trials were partially encoded, the small group partially encoded a total number of 144 instances, which were all target items. The large set group completed 540 trials. If the large set group took longer to learn what information was redundant and thus only started to partially encode during Block 15 and all further trials were partially encoded, the large group partially encoded a total number of 270 instances, but only 90 partial instances were target items. If this occurred,
what would be the effect of having more partially encoded target instances for the small set group versus the large set group, but more total partially encoded instances for the large set group than for the small set group?

One possible effect of more total trials partially encoded for the large set group than for the small set group may be that persisting to retrieve partially encoded instances during transfer lasted longer for the large set group than for the small set group. Indeed, persistence in retrieving partial instances is supported by the ACT-R framework of cognitive skill acquisition (e.g., Anderson & Lebiere, 1998).

ACT-R is a strength-based framework that describes responding to an item of a cognitive skill as the interaction of declarative memory with procedural memory. Most important here is procedural memory. Procedural memory in ACT-R is the mechanism for item-general, rule-based processing, which is instantiated as a series of productions. Generally, productions compete with each other to be activated. In part, the likelihood of activation is higher for a production that has a history of successfully leading to a correct response versus an incorrect response. With practice, multiple productions may be consolidated into a single production. For example, a possible consolidated production for partial retrieval of a true AA item is:

$\text{IF}$ the goal is to do an alphabet arithmetic problem

$\text{AND}$ there is a fact in memory stating that letter1 = TRUE

$\text{THEN}$ retrieve fact from declarative memory.

Importantly, if this production leads to a correct response, it accrues strength. Strength determines both the likelihood of selection and the speed of the production on subsequent
Regarding Experiment 1, the partial-item production strength would be greater at the end of practice for the large set group versus the small set group because the production was successfully completed during practice more times for the large set group than for the small set group. Thus, the likelihood of using the production in Transfer Block 1 would be higher for the large set group versus the small set group. Conversely, the likelihood of using the initial sequence of productions for solving an AA problem would be less for the large set group versus the small set group.

If persistence in retrieving partial instances lasted longer for the large set group than for the small set group, all Experiment 1 transfer results may be accounted for. Specifically, longer persistence retrieving partial instances for the large versus small set group explains why novel item accuracy was worse for the large set group than for the small set group. Retrieving partial instances for novel items was more likely for the large set group versus the small set group. In addition, the persistence account explains why response times for target items during Transfer Block 1 were faster for the large set group versus the small set group. Speed of responding is faster when partially attending to stimuli versus fully attending to stimuli. Finally, the persistence account explains why addend slope for target items during transfer was lower for the large set group versus the small set group. Reversion to rule-based processing was less likely for the large set group versus the small set group. In sum, learning to reduce information may have been faster for the small set group versus the large set group. However, persistence in retrieving partial instances during transfer may have been shorter for the small set group versus the large set group.
INTRODUCTION EXPERIMENT 2

Although Experiment 1 results supported the idea that the practice context manipulation of set size influences gains made during practice and has consequences for later performance, not all of Experiment 1’s results were of the expected pattern. For example, novel item accuracy during Transfer Block 1 was unexpectedly worse for the large set group versus the small set group. Given this unexpected result was not consistent with my original assumptions of the set size effect, Experiment 1’s support for the hybrid theory is questionable. However, predictions of Experiment 1 did not account for the influence of practicing filler items. Partial encoding of filler items during practice may have influenced persistence in retrieving partial target instances during transfer. Specifically, persistence retrieving partial target instances during transfer was shorter for the small set group versus the large set group. Although the persistence account explains results of Experiment 1, the account remains unverified.

Experiment 2 was designed to replicate the results of Experiment 1 and afforded an investigation of the alternative persistence account for the list length effect and thus the hybrid theory. Experiment 2 shared the same design as Experiment 1 except a third group (small set plus filler group) was added that completed the same number of trials as the large set group. In order for the trials to be equated for small set plus filler group and large set group, the small set plus filler group practiced the filler set of 12 items first in practice, before practicing the six target items.
Regarding Experiment 2 results between the small versus large set groups, results of Experiment 2 were expected to be the same as Experiment 1. Regarding results comparing the small set plus filler group to the small and large set groups, predictions of Experiment 2 were made on the basis of the alternative account for the list length effect.

Recall that the alternative persistence account is the same assumptions as the original assumptions about set size except for the assumption that the total number of trials partially encoded influences persistence of retrieving partially encoded target instances during transfer. If the alternative persistence account is correct, the rate of learning redundant information during practice would be greater for the small set plus filler group than for the large set group because learning redundant information is easier in practicing a small set of items versus a large set of items. However, the rate of learning redundant information would not be significantly different between the small set plus filler group and the small set group because each groups’ practice sets contain the same number of items. In addition, given that learning redundant information during practice is an item-general process, once the small set plus filler group learns to attend and thus encode partial instances during practice of the filler items, they will continue to partially attend to and thus partially encode items of the subsequently presented target set. Taken together, three assumptions concerning differences between the small set plus filler group, the small set group, and the large set group were made. First, the number of target items partially encoded during practice would be greater for the small set plus filler group than for the large set group. Second, the number of target items partially encoded during practice would be greater for the small set plus filler group than the small set group
because the small set plus filler group already learned to partially attend/encode stimuli before practicing the target items. Third, the total number of practice trials partially encoded would be greater for the small set plus filler group than for the large set group. Thus, although rate of learning redundant information is not significantly different between the small set plus filler group and the small set group, persistence retrieving partially encoded instances during transfer will be longer for the small set plus filler group than for the small set group. In short, the small set plus filler group’s performance during transfer will be more like the large set group than the small set group. Predictions and results regarding the alternative account will be discussed at the end of Experiment 2’s results section.

Experiment 2 affords investigation of another alternative explanation of the Experiment 1’s practice results. Experiment 1’s practice results may be accounted for by mental fatigue. If mental fatigue was greater for the large set group than for the small set group. Response times for target items at the end of practice may have been faster for the small set group than for the large set group because of the larger number of practice trials for the large set group (540) versus the small set group (180). If mental fatigue was a factor, retrieval from memory at the end of practice would be slower for the large set group than for the small set group. However, if mental fatigue was higher for the large set group than for the small set group in Experiment 1, it is unclear why response times for target items during transfer were faster for the large set group versus the small set group. If mental fatigue completely accounts for practice differences in Experiment 1, response times for target items at the end of practice will be significantly slower for the small set
plus filler group than for the small set group in Experiment 2 because more trials are practiced in the small set plus filler group versus the small set group.
METHOD EXPERIMENT 2

Participants and Design.

Ninety undergraduates from Kent State University participated for research credit in an introductory psychology class. Participants were randomly assigned to one of three practice groups (small set, small set plus filler, or large set), defined by the number of items practiced in target item blocks (small versus large) and the presence of filler items.

Materials and Procedure.

Stimuli consisted of the same 24 alphabet arithmetic items in Experiment 1. For the small set and large set groups, the procedure was identical to Experiment 1. For the small set plus filler group, in the practice session, filler items were broken up into two sets of six items. Each set had three true items and three false items. For true and false items, one item was of addend size 2, one was of addend size 3, and one was of addend size 4.

After completing warm-up trials, the small set plus filler group practiced one set of filler items, repeated for 30 blocks, which was immediately followed by 30 practice blocks of the second filler set. Immediately after the two filler sets were practiced, participants practiced the target set for 30 blocks in the same manner as Experiment 1. For all groups, transfer was identical to Experiment 1. After completing practice and transfer, participants verified 18 multiplication problems. Here, for each trial, a multiplication problem was presented with an answer (e.g., 24 X 12 = 288) and two response buttons
(TRUE versus FALSE), which were directly below the multiplication problem. Half of the problems were true and half were false. Verification of multiplication problems was used as another measure of mental fatigue.

**Results.**

One participant performed below 75% accuracy and was dropped from analyses. Except where otherwise noted, analyses of response times were conducted on correct response trials only (excluding response times less than 50ms and greater than 9000ms). Outliers were treated the same in Experiment 2 as in Experiment 1.

**Individual differences.** Regarding fatigue, a one-way ANOVA on participants’ mental fatigue ratings resulted in a significant effect of group, $F(2,84) = 4.24, \text{MSE} = 6.05, p = .018$. Tukey’s HSD post hoc test resulted in significantly lower mental fatigue ratings for the small set group ($M = 4.77, \text{SE} = 0.44$) than the large set group ($M = 6.57, \text{SE} = 0.52$). Mental fatigue ratings for the small set plus filler group ($M = 5.17, \text{SE} = 0.42$) were not significantly different from either the small set group or the large set group. In addition, one-way ANOVAs on response times and accuracy for multiplication problems resulted in no significant differences between groups, all $Fs < 1.67, all ps > .195$. Together, these results strongly suggest that fatigue was not a causal factor for performance difference between groups in Experiment 2. Indeed, the pattern of Experiment 2 results was not different between tests with fatigue measures as a covariate versus tests without fatigue measures as a covariate. Thus, for simplicity, reported results are from tests without fatigue measures as a covariate.
Regarding primary language, six participants reported that English was a secondary language. Results of mixed factor ANOVAs for target response times during practice, target response times during test, target items’ addend slope during test, and overall accuracy during test found no significant differences between primary and secondary English speakers, all $F$s < 1.10, all $p$s > .298. Thus, all results reported below include both primary and secondary English speakers.

**Practice session.** For each participant, response times for the six target items were averaged for each block of practice and for each block of transfer. Similarly, for transfer blocks containing novel items, response times for the six novel items were averaged. Mean response time data for target and novel items are displayed in Figure 4.
Regarding Prediction 1, inspection of Figure 4 reveals that target item response times during practice were significantly faster for the small set group versus the large set group. A 2 (group: small set versus large set) X 30 (practice block) mixed-factor ANOVA for target item response times resulted in significant main effects of group and block, and a significant group X block interaction, $F(1,57) = 15.19$, MSE = 6515950, $p < .001$, $F(29,1653) = 141.96$, MSE = 272391, $p < .001$, and $F(29,1653) = 3.54$, MSE = 272391, $p < .001$ respectively. Furthermore, an independent-samples t-test for target item response times in the last block of practice confirmed that response times were significantly faster for the small set group versus the large set group at the end of
practice, $t(58) = 3.63, SE = 113, p = .001$, which is consistent with the predictions of the hybrid theory. Importantly, the results are inconsistent with instance theory, which would predict no significant response time differences during practice between small versus large set groups.

Concerning the alternative account of fatigue during practice, if 540 trials produce a fatigue effect and the fatigue effect is the only factor influencing response time differences between groups, target item response times at the end of practice will not be significantly different between the small set plus filler group and the large set group during practice. However, as can be seen in Figure 4, response times for target items at the end of practice were faster for the small set plus filler group than for the large set group. An independent-samples $t$-test for target items in the last block of practice resulted in significantly faster response times for the small set plus filler group than the large set group, $t(56) = 5.82, SE = 106, p < .001$, which is inconsistent with the idea that fatigue was a causal factor for performance differences between groups during practice. Given that results strongly suggest that fatigue was not a factor in Experiment 2, fatigue effect on performance during transfer will not be discussed.

*Transfer.* Percentages of target and novel items correct for each participant were averaged for each block of transfer. Mean percentage of target and novel trials correct for each block of transfer and each practice group are displayed in Figure 5.
Regarding Prediction 2, inspection of Figure 5 reveals that novel item accuracy during the first block of transfer was higher for the small set group than for the large set group. An independent samples $t$-test for percentages of novel items correct in Transfer Block 1 resulted in significantly higher novel item accuracy for the small set group ($M = 84.4\%, SE = 3.2$) versus the large set group ($M = 65.5\%, SE = 3.4$), $t(58) = 4.06$, $SE = 4.65$, $p < .001$, which inconsistent with instance theory.

Regarding Prediction 3, a 2 (group: small set versus large set) X 10 (transfer block) mixed factor ANOVA for target item response times resulted in a significant main effect of block, $F(9,522) = 6.85$, MSE = 119693, $p < .001$. Although the main effect of group was not significant [$F(1,58) = 0.79$], the prediction of the hybrid theory was most
concerned with target item response times at the beginning of transfer. Thus, an independent-sample $t$-test on target item response times was during the first transfer block was conducted. Response times for target items were not significantly different between the small set group and the large set group, $t(58) = 0.70$. These results are inconsistent with the hybrid theory and the idea that reversion back to using the rule-based process was higher for the small set group than for the large set group. It is unclear why response times were slower for the small set group versus the large set group in Experiment 1 but not different in Experiment 2 because Experiment 1 methods for these groups were the same as Experiment 2. Perhaps the difference between small set groups of Experiment 1 versus 2 reflects sampling error.

![Figure 6. Experiment 2 addend slopes of trials corrected for target items as a function of set size group and mini blocks of trials during transfer. Error bars equal standard error.](image)

Figure 6. Experiment 2 addend slopes of trials correct for target items as a function of set size group and mini blocks of trials during transfer. Error bars equal standard error.
For Prediction 4, a 2 (group: small set versus large set) X 5 (mini block) mixed factor ANOVA for addend slope resulted in a significant main effect of group, $F(1,58) = 14.59$, MSE = 228676, $p < .001$, which is consistent with the hybrid theory. However, an independent samples $t$-test on target item addend slopes during the first mini block of transfer resulted in no significant target item addend slope differences between the small set group ($M = 244\text{ms/addend}, \text{SE} = 99$) than for the large set group ($M = 107\text{ms/addend}, \text{SE} = 55$), $t(58) = 1.21, p = .233$, which was possibly due to the small number of items of each addend size and/or the large variability inherent in addend slope data.

Regarding the alternative persistence account, results at the end of practice were consistent with the idea that the number of partially encoded target instances during practice was greater for the small set plus filler group versus both the large set group and the small set group. Response times for target items at the end of practice were significantly faster for the small set plus filler group versus the large set group, $t(56) = 5.82, \text{SE} = 106, p < .001$. Response times for target items were significantly faster for the small set plus filler group versus the small set group, $t(58) = 3.75, \text{SE} = 56, p < .001$. Furthermore, this second result is consistent with the idea that participants learned to partially encode during practice of filler sets and that these item-general gains then carried over to practice of the target items. Also, given that more total trials were practiced for the small set plus filler and large set groups versus the small set group, the alternative account assumes that more total trials were partially encoded during practice for the small set plus filler group and the large set group than for the small set group.
If true, the alternative account assumes that persisting to retrieve partial instances during transfer would last longer for the small set plus filler and large set groups than for the small set group, which would cause accuracy for novel items to be worse for the small set plus filler and large set groups versus the small set group. Indeed, during Transfer Block 1, accuracy for novel items was significantly lower for the small set plus filler group than for the small set group, $t(58) = 2.66$, SE = 5, $p = .01$. And, accuracy for novel items was not significantly different between the small set plus filler group and the large set group, $t(56) = 0.97$. Together, these results are consistent with the alternative persistence account.

During Transfer Block 1, although response time for target items was neither expected to, nor differed between the small set plus filler group and the large set group [$t(56) = 0.68$], response time for the target items was expected to be significantly faster for the small set plus filler group versus the small set group. However, $t$-test results were not significantly different between the small set plus filler group versus the small set group, $t(56) = 0.55$, which is inconsistent with the alternative account.

During Transfer Mini Block 1, although addend slope for target items was neither expected to, nor differed between the small set plus filler group and large set group [$t(56) = 0.85$], addend slope for target items was expected to be significantly lower for the small set plus filler group versus the small set group. A $t$-test resulted in significantly lower addend slope for the small set plus filler group (M = 19 msec/addend, SE = 86) versus the small set group (M = 245 ms/addend, SE = 99), $t(58) = 1.71$, SE = 132, $p = .045$, which is consistent with the alternative account.
INTRODUCTION EXPERIMENT 3

Results of Experiments 1 and 2 suggest that practice context influences what is encoded in an instance and thus what information is available for later retrieval. Furthermore, the overall pattern of results of Experiments 1 and 2 are consistent with the hybrid theory’s predictions, which were made on the basis of assumptions about the set size effect. Experiment 2 results also ruled out the possibility that performance differences were due to fatigue.

Experiment 3 was conducted to extend the result of Experiments 1 and 2 beyond implicit, automatic tasks to an explicit recognition task in order to further test the idea that individuals learn to partially attend to and thus partially encode stimuli during practice. Experiment 3 methods were the same as Experiment 1 except the transfer phase involved a speeded recognition test. For the recognition test, all participants were presented with the set of target items, and a set of foil items. Foil items were matched to target items and shared terms with a given target item. By matching terms of foil items to target items, participants that attended to and thus partially encoded target items during practice will make retrieval errors during the speeded recognition task. Specifically, according to the hybrid theory’s account of the set size effect, partially attending to foil items in the speeded recognition task will cause partially encoded target instances to be retrieved and thus foil items will be incorrectly recognized as target items.
Given that the practice session of Experiment 3 is the same as Experiment 1, Prediction 1 is the same as in Experiment 1. Of most importance here are the predictions regarding the speeded recognition task. However, before getting to these predictions, note that correctly responding in the speeded recognition task may only be accomplished by correct use of the retrieval process. Using the rule-based process to generate the correct answer to an AA item is not useful to recollecting that you have seen the item before. Thus, regarding speed of recognition, if the retrieval process is faster for partial instance retrieval versus whole instance retrieval, response times for false alarms (recognizing foils as previously seen) will be faster for the small set group versus the large set group. This prediction reflects the idea that during foil trials more partial target instances are retrieved for the small set group versus the large set group. Regarding recognition accuracy, if more target items were partially encoded during the practice phase for the small set group versus the large set group, accuracy (corrected recognition) during the recognition task will be worse for the small set group versus the large set group. This prediction reflects the idea that the small set group will have few whole-stimulus instances to accurately respond and will incorrectly retrieve partial target instances for foils.
METHOD EXPERIMENT 3

Participants and Design.

Eighty-six undergraduates from Kent State University participated for research credit in an introductory psychology class. Participants were randomly assigned to one of two practice groups (small set or large set), defined by the number of items practiced.

Materials and Procedure.

Experimental stimuli were the same AA items from Experiments 1 and 2. Procedures were identical to Experiment 1 small set group and large set group, with the following exceptions. During the speeded recognition test, participants responded to six target items presented three times each and 18 foil items presented once each. Each target item had three different types of foil items matched to it (see Table A2 in the Appendix). Type 1 foil items shared the first two terms with a given target item (Foil 1). Type 2 foil items shared the first term and the response with a given target item (Foil 2). Type 3 foil items shared first and third terms with a given target item (Foil 3).

Given that each target item was matched with three foil items during the speeded recognition test, there were six YES responses and 18 NO responses total between these items. In order to have an equal number of YES and NO responses during the speeded recognition test, the target items were repeated three times each in blocks. In each block of transfer, target items were presented with only one type of matched foil. Together, the
speeded recognition test consisted of 36 trials, with half of the trials correct responses being YES and half of the trials correct responses being NO.

Immediately after completing the practice phase, participants were given instructions for the speeded recognition test. Participants were instructed that in the next part of the experiment their task was to judge whether or not they saw each item in the practice phase. Furthermore, participants were instructed that they needed to answer each trial as quickly and as accurately as possible. Participants proceeded directly to the recognition test after reading the instructions.

On each speeded recognition trial, an orientation stimulus (***) was presented for 500 ms, followed by an AA item, the question and two response buttons (‘‘YES’’ and ‘‘NO’’) appearing below the AA item. After participants clicked on a response button, the stimulus and the response buttons disappeared. If the response was incorrect, a red ‘‘ERROR’’ message was presented for 1000 ms, followed by a button labeled ‘‘next’’. If the participant’s response was correct, the ‘‘next’’ button was immediately presented. The participant then clicked on the ‘‘next’’ button to present the orientation stimulus for the next trial, followed by the next AA item, and so on. Response time was recorded as the time between stimulus onset and clicking one of the two response buttons. Feedback concerning average speed and accuracy was presented after every 48 trials. Participants were also given the opportunity to take a small break during the presentation of feedback. After completing the speeded recognition task, participants reported mental fatigue and indicated if English was their primary language as in Experiments 1 and 2. Participants then solved multiplication problems in the same way as Experiment 2.
Results.

Two participants performed below 75% accuracy and were dropped from analyses. Except where otherwise noted, analyses of response times were conducted on correct response trials only (excluding response times less than 50ms and greater than 9000ms). Outliers were treated the same as in Experiments 1 and 2.

Individual differences. Regarding fatigue, ratings of fatigue were significantly greater for the large set group (M = 5.55, SE = 0.32) than for the small set group (M = 3.43, SE = 0.30), $t(84) = 4.81, p < .001$. However, interpretation of inferential results was not different between tests with fatigue as a covariate and tests without fatigue as a covariate. Thus, for simplicity, reported results are from tests without fatigue reports as a covariate. Regarding verification of multiplication problems, independent samples t-tests on response times and accuracy resulted in no significant differences between groups, all $t$s < 1.15, and all $p$s > .255. Thus, neither response times nor accuracy of multiplication verification was used as a covariate in the following analyses.

Regarding primary language, nine participants reported that English was a secondary language. A 2(English: Primary versus Secondary) X 30 (block) mixed factor ANOVA for target response times during practice resulted in a significant English X block interaction, $F(29,2407) = 1.57, \text{MSE} = 294110, p = .027$. However, given that the nine participants were nearly evenly divided between the small set group (five participants) and the large set group (4 participants), reported analyses include both primary and secondary English speakers.
**Practice session.** For each participant, response times for the six target items were averaged for each block of practice separated by small and large set groups. Mean response time data for target items are displayed in Figure 7.

![Figure 7](image)

**Figure 7.** Experiment 3 mean response times as a function of set size group and blocks of trials. Error bars equal standard error.

Regarding Prediction 1, a 2 (group: small set versus large set) X 30 (practice block) mixed-factor ANOVA for target item response times resulted in significant main effects of group and block, and a significant group X block interaction, $F(1,83) = 20.53$, $\text{MSE} = 8383671, p < .001$, $F(29,2407) = 195.94$, $\text{MSE} = 280305, p < .001$, and $F(29,2407) = 5.74$, $\text{MSE} = 280305, p < .001$ respectively. Furthermore, a one-tailed independent samples $t$-test for target item response times in the last block of practice confirmed that response times were significantly faster for the small set group than for the large set group at the end of practice, $t(58) = 1.97$, $\text{SE} = 101, p = .026$, which is
consistent with the predictions of the hybrid theory and the idea that small set group was encoding partial instances.

*Speeded recognition test.* For each participant, response times for correct recognition of target items (Hit), incorrect recognition of target items (Miss), correct recognition of foils (Correct Rejection), and incorrect recognition of foils (False Alarm) were averaged across each block of the speeded recognition task, separated by group. Response time means for Hits, Misses, Correct Rejections, and False alarms separated by group are displayed in Figure 8.

![Figure 8](image)

**Figure 8.** Experiment 3 response times for speeded recognition task as a function of Hit, Miss, Correct Rejection, and False Alarm. Response times are further separated by set size group. Values are mean response times in milliseconds. Error bars equal standard error.

Regarding speed of responding, response times for false alarms were predicted to be faster for the small set group versus the large set group, which occurred (see Figure 8).
However, separate independent-samples $t$-tests on response times for Hit, Miss, Correct Rejection, and False Alarm resulted in response times being significantly faster for the small set group than for the large set group only for Correct Rejections, $t(84) = 2.64$, SE = 152, $p = .010$ (all other $ts < 0.63$, all other $ps > .52$). Thus, the overall pattern of results is inconsistent with the idea that retrieval of partial target instances for foils was more likely for the small set versus large set group. Indeed, if this reasoning was correct, response times for correct rejection should have been slower for the small set group versus the large set group because the number of whole instances available for retrieval was thought to be less for the small set group versus large set group.

For each participant, accuracy was calculated as the number of hits – the number of false alarms (corrected recognition) in each block. Then, for each participant, accuracy was averaged across the three blocks of the speeded recognition task. Group means of Hits, False Alarms and Corrected Recognition are displayed in Table 1.

Table 1. Experiment 3 mean proportion of Hits, False Alarms, and Corrected Recognition as a function of set size group.

<table>
<thead>
<tr>
<th>Group</th>
<th>Hits M (SE)</th>
<th>False Alarms M (SE)</th>
<th>Corrected Recognition M (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Set</td>
<td>.90 (.02)</td>
<td>.18 (.02)</td>
<td>.71 (.04)</td>
</tr>
<tr>
<td>Large Set</td>
<td>.87 (.02)</td>
<td>.29 (.03)</td>
<td>.57 (.04)</td>
</tr>
</tbody>
</table>

*Note. M = mean, SE = standard error.*

Regarding accuracy, if the likelihood of retrieving partial target instances for foil trials was higher for the small set group versus large set group, accuracy would be lower for the small set group than for the large set group. As can be seen in Table 1, this was
not the case. An independent samples $t$-test on corrected rejection resulted in accuracy being significantly higher for the small set group than the large set group, $t(84) = 2.41$, $SE = 0.33$, $p = .018$.

Although the patterns of results for the speeded recognition task are inconsistent with predictions, note that so far the results presented were collapsed across three Transfer Blocks. Given the possibility that aggregating across blocks may have created a different pattern of results than what occurred at the beginning of the speeded recognition task, analysis of data in Recognition Block 1 was conducted. However, all patterns of results for Block 1 were the same as aggregated results. Only response time for correct rejection was significantly faster for the small set group versus the large set group, $t(84) = 2.30$, $SE = 177$, $p = .024$. And, corrected recognition was higher for the small set group versus the large set group, $t(84) = 2.70$, $SE = .07$, $p = .008$.

In sum, although the practice results were consistent with the hybrid theory, recognition results were inconsistent with the hybrid theory. Indeed, the accuracy results are consistent with the idea that fewer instances were partially encoded during practice for the small versus large set group. This unexpected pattern of results will be considered further in the General Discussion.
GENERAL DISCUSSION

Empirically, the goal of the current study was to extend the memory-based automaticity literature in regards to what is encoded during practice, by addressing two specific questions: Does practice context influence what information is encoded? If so, does differential encoding during practice have consequences on later performance? In this regard, results supported the idea that practice context influences what information is encoded and expanded previous research by showing that set size affects the gains made during practice and the durability of gains after practice. Key results consistent with set size influencing practice gains occurred in all three experiments. In all three experiments, response times for target items at the end of practice were significantly faster for the small set group versus the large set group. Furthermore, key results supported the idea that differential practice gains have consequences on later performance. In Experiments 1 and 2, accuracy for novel items was worse for the large set group than the small set group, and response times for target items were slower for the small set group than for the large set group (although not significant in Experiment 2). In Experiments 1 and 2, addend slope for target items tended to be lower for the large set group than for the small set group. Although addend slope differences were not significant, this may have been due to insufficient power to find an effect. In Experiment 3, response times for correct rejections were faster and accuracy was higher for the small set group versus the large set group.
Implications for Memory-Based Processing Theories of Automaticity.

Importantly, MBP theories of automaticity (e.g., Logan, 1988; Palmeri, 1997; Rickard, 1997) cannot account for the current results. Although MBP theories describe practice gains as a shift from rule-based processing to retrieval, no MBP theories have a mechanism to account for practice context effects. Thus, the hybrid theory proposed here was an attempt to extend memory-based processing theories to account for practice context effects.

The hybrid theory is an extension of instance theory that incorporates the rule-based practice gain mechanism of information reduction theory (Haider & Frensch, 1996). Briefly, the key mechanism of the hybrid theory that accounts for the practice context effect was learning to ignore redundant information. According to the hybrid theory, once individuals learn to ignore task redundant information, they will only attend to task relevant stimulus information. Thus, task relevant stimulus information will be the only information encoded into memory and available for retrieval later.

In order to test the hybrid theory, the current experiments manipulated practice context through comparing small versus large set groups. The assumption being that the rate of learning redundant information and thus attending to and encoding partial instances would be faster for the small set group versus the large set group. Set size was chosen to manipulate practice context here because of previous explicit memory studies of the list length effect, which found that retrieval from memory was more likely and faster when studying a small list versus a large list. Presumably, the reason for the list length effect is that short lists afford greater item-specific processing than long lists.
whereas long lists afford more contextual processing than short lists. One way to specify greater item-specific processing is through information reduction, which is consistent with the idea of focusing attention on specific items rather than on the relationship between items.

Although results for practice and transfer differed between the two groups in all experiments, differences were not always in the direction of predictions made on the basis of previous studies (e.g., Experiment 3 accuracy was higher for the small set group versus the large set group). Given these unexpected results, more investigation of the set size effect in implicit memory is needed. Furthermore, given that further investigation of the set size effect is required, further investigation is also required to definitively evaluate the hybrid theory. Two possible ways to further evaluate the effect are discussed in the Future Directions section. Here, I will focus the rest of this section on the possible nature of the set size effect.

Generally, the set size effect may reflect partial encoding or it may reflect full encoding with the addition of some other sort of information. Regarding partial encoding, the original idea was that the number of target items partially encoding would be higher for the small set group than for the large set group, and thus accuracy in Experiment 3 would have been worse for the small set group than for the large set group. In contrast, accuracy was higher for the small versus large set group. This result raises the possibility that the number of partially encoded instances was lower for the small set versus large set group. However, this possibility cannot account for other outcomes, such as faster response times during practice for the small set versus large set group.
Although the original idea cannot account for all results of Experiment 3, the alternative persistence account may. Recall that the persistence account still assumes that more partial target instances are encoded during practice for the small set group versus the large set group. However, the persistence account assumes that retrieval of partially encoded target instances during transfer would last for a shorter amount of time for the small set group versus the large set group because of the influence of filler items on the large set group. To this point, description of the alternative account has focused on the idea that both groups may learn to encode partial instances during practice. However, now the focus will turn to the idea that both groups also encode full instances during practice. For example, in Experiment 3, both groups presumably encoded full instances at the beginning of practice. At the beginning of practice neither group would have learned what information to ignore. Thus, during the speeded recognition task both groups had the required information (fully encoded instances) to respond correctly. If true, and if persisting to retrieve partial instances lasted shorter after practice for the small set group versus the large set group, it is possible that the likelihood of retrieving full instances would be higher for the small set versus large set group. Experiment 3 accuracy results in the speeded recognition task are consistent with this idea. Regarding speed of responses during the speeded recognition task, the persistence account’s explanation fares less well than the explanation for accuracy. Consistent with the persistence account, response times for Hits and Correct Rejections were faster for the small set group versus the large set group because the likelihood of first retrieving whole target instances was higher for the small set group versus the large set group. Relative to the small set group, the large
set group was more likely to first retrieve a partial target instance, fail to respond, and then retrieve whole target instances to make a Hit or Correct Rejection. However, the persistence account would predict that response times for Misses and False Alarms would be slower for the small set group versus the large set group because the retrieval process is slower for whole instances versus partial instances.

Of course, the set size effect may also reflect encoding full instances along with some sort of other information. One possible explanation here is encoding variability theory (e.g., D’Agostino & DeRemer, 1973; Glenberg, 1979; Melton, 1970). In general, encoding variability theory assumes that individuals encode both the stimulus and the context in which the stimulus is embedded. Retrieval speed is faster and retrieval likelihood is higher when the context at time of retrieval matches the context at encoding. Regarding the current experiments, context may be thought of as the item in the practice trial immediately preceding a target trial. If so, the likelihood of the context for a target item being the same item on each presentation during practice would be higher for the small set group versus the large set group. Thus, context varied less for small set versus the large set group, which accounts for response times at the end of practice being faster for the small set versus large set group.

Although contextual variability theory does well in accounting for the current experiments’ practice results, it may not account for all transfer results of Experiments 1 and 2. Relative to practice, contextual variability is increased during transfer for the small set group but decreased for the large set group in Experiments 1 and 2. For the short set group, a specific target item may have been preceded by five items during practice and
eleven items during transfer. For the large set group, a specific target items may have been preceded by 17 items during practice and 11 items during transfer. Thus, if change in contextual variability matters, target item accuracy during transfer would be worse and response time would be slower for the small set group versus the large set group.

Although in Experiments 1 and 2 response time for target items during transfer was slower for small versus large items, accuracy for target items did not differ between the groups. However, accuracy of target items was near ceiling for both groups. Importantly though, the contextual variability account provides no explanation for differences in novel item accuracy during transfer. Contextual variability requires previous practice and by definition novel items have not been practiced.

Another explanation for the current results that does not require partial instances is interference (e.g., see Wilkins & Rawson, under review). Although only generally specified, the idea here is that memory for items interfere with the retrieval of a specific item during practice and test. Furthermore, interference of retrieval for a specific item is higher when the specific item was previously practiced in a low interference context versus a high interference context. Thus, the interference hypothesis presents a straightforward account of the practice results in the current experiments. Response times for target items at the end of practice were faster for the small set group versus the large set group because interference was less during practice for the small versus large set group. The interference hypothesis also does well explaining the transfer results of Experiments 1 and 2. During Transfer Block 1, memory for items previously practiced will interfere with novel items. Thus, according to this hypothesis, novel item response
times would be faster and accuracy would be greater for the small set group versus the large set group because fewer items are in memory to interfere with novel items for the small set group than for the large set group. Although accuracy results for novel items in Transfer Block 1 were consistent with the interference hypothesis for both Experiments 1 and 2, the pattern of response time results for novel items were the exact opposite. Response times for novel items were descriptively slower for the small set versus the large set group. However, note that to correctly respond to novel items in Transfer Block 1, responses must be generated by rule-based processing rather than retrieval-based processing. Thus, the pattern of response time results for novel items may reflect that the amount of rule-based practice gains was less for the small set group than for the large set group. Indeed, Wilkins and Rawson (2010) found more rule-based practice gains for large set items versus small set items. If this occurred during practice, response times for novel items in Transfer Block 1 may reflect greater interference for the large versus small set group but interference is offset by greater rule-based gains during practice for the large versus small set group. Further investigation is required to pull apart the two effects. One way to do this would be to incorporate novel items throughout practice that are not foils of target items. Having novel items during practice would give a pure measure of the speed of rule-based processing throughout practice. However, if the hybrid theory is correct and partial instances are encoded, the addition of novel items during practice may hamper the learning of redundant information.

In sum, the current research added to the small number of previous memory-based processing studies that suggest that practice context influences what information is
encoded during practice. Current memory-based processing theories have no mechanism to account for the present findings. Thus, memory-based processing theories need to be further instantiated to account for practice context effects. Although the current paper proposed the hybrid theory to account for practice context effects, the hybrid theory requires further investigation to be verified. Verification of the hybrid theory was not possible with the current experiments because the nature of the set size effect in an implicit automaticity task is not well understood. For this reason, further investigation of the set size effect is also needed.
FUTURE DIRECTIONS

The current research is one of few studies that suggest that practice context influences what information is encoded during practice and that differential encoding during practice has consequences for later performance. Also, the current research only investigated one type of practice context manipulation, set size. The current research makes apparent that the set size effect is not well understood. Given the lack of understanding the set size effect, the hybrid theory was not able to be definitively evaluated. Thus, future research is needed in many different areas, including other practice context manipulations, to definitively evaluating the hybrid theory and the idea of partial encoding. In order to keep the Future Directions section relatively brief, I will provide an example of two methods for investigating if partial instances are encoded.

One way to examine if partial instances are encoded is to use similar methods as Experiments 1 and 2, but change the transfer task to include novel items that are only partial representations of previously practiced target items. For example, if during practice the target item $A + 2 = C$ was repeatedly presented, the partial item during transfer would be $(A \quad )$. If during practice more partial instances of target items are encoded for the small set group than for the large set group, partial item accuracy would be higher and response time during transfer would be faster for the small set group versus the large set group. However, if partial instances are not encoded, partial item accuracy and response times would not significantly differ between small versus large set groups.
Another way to examine if partial instance are encoded during practice would be to use a moving window. This proposed experiment would have the same methods as Experiment 1, with the addition adding the moving window component and vocal responding. The key difference here from Experiment 1 would be that not all terms of a presented stimulus would be initially presented. Rather, only the first term of an item would be presented at the beginning of a trial (e.g., \( A + \ = \)). Participants would be able to reveal the next term of the item by pressing the spacebar. So, if the space bar was pressed once, the item would look like \( A + 2 = \). Participants would vocally respond TRUE or FALSE at any time during a trial, which would be recorded and measured with a computer microphone. Thus, if participants learned to ignore redundant information they would respond TRUE or FALSE without ever pressing the spacebar. The number of spacebar presses during practice would be predicted to be less for the small set group versus the large set group because learning that the hidden information is redundant would be faster for the small set group than for the large set group. Research such as this could offer valuable insight towards how to develop training context that afford the largest gains during practice that are the most durable.
REFERENCES


APPENDIX

Table A1. Materials for Experiments 1 and 2.

<table>
<thead>
<tr>
<th>Target</th>
<th>Novel</th>
<th>Filler 1</th>
<th>Filler 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C + 2 = E</td>
<td>C + 2 = F</td>
<td>X + 2 = Z</td>
<td>D + 2 = F</td>
</tr>
<tr>
<td>H + 3 = K</td>
<td>H + 3 = L</td>
<td>A + 3 = D</td>
<td>T + 3 = W</td>
</tr>
<tr>
<td>L + 4 = P</td>
<td>L + 4 = Q</td>
<td>Q + 4 = U</td>
<td>F + 4 = J</td>
</tr>
<tr>
<td>V + 2 = Y</td>
<td>V + 2 = X</td>
<td>J + 2 = M</td>
<td>U + 2 = X</td>
</tr>
<tr>
<td>O + 3 = S</td>
<td>O + 3 = R</td>
<td>R + 3 = V</td>
<td>K + 3 = O</td>
</tr>
<tr>
<td>I + 4 = N</td>
<td>I + 4 = M</td>
<td>G + 4 = L</td>
<td>B + 4 = G</td>
</tr>
</tbody>
</table>

Table A2. Example of materials for Experiment 3 speeded recognition test.

<table>
<thead>
<tr>
<th>Type</th>
<th>Item</th>
<th>Truth Value</th>
<th>Correct Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>C + 2 = E</td>
<td>True</td>
<td>Yes</td>
</tr>
<tr>
<td>Foil 1</td>
<td>C + 2 = F</td>
<td>False</td>
<td>No</td>
</tr>
<tr>
<td>Foil 2</td>
<td>C + 3 = F</td>
<td>True</td>
<td>No</td>
</tr>
<tr>
<td>Foil 3</td>
<td>C + 3 = E</td>
<td>False</td>
<td>No</td>
</tr>
</tbody>
</table>