WHEN MORE TIME HURTS PERFORMANCE:
A TEMPORAL ANALYSIS OF ERRORS IN EVENT COUNTING

A Thesis in
Psychology

by

Daniel N. Cassenti

© 2004 Daniel N. Cassenti

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2004
The thesis of Daniel N. Cassenti was reviewed and approved* by the following:

Richard A. Carlson  
Professor of Psychology  
Thesis Adviser  
Chair of Committee

Judith F. Kroll  
Liberal Arts Research Professor of Psychology and Linguistics

Daniel J. Weiss  
Assistant Professor of Psychology

Frank E. Ritter  
Associate Professor of Information Sciences and Technology

Kevin R. Murphy  
Professor of Psychology  
Head of the Department of Psychology

*Signatures are on file in the Graduate School
Abstract

The speed-accuracy tradeoff suggests that less time to process information leads to less accurate performance. However, when individuals are forced to process information more slowly this also may lead to less accurate performance. Four empirical studies based on a counting task investigated the theory that more time hurts performance and found consistent, confirmatory support. The findings indicate that although forgetting may negatively affect counting, mental rehearsal, a strategy aimed at preventing forgetting may actually increase the chance of making an error. Findings indicate that distracting information may also negatively influence performance. Working memory span, number representation, and other issues are explored and discussed.
# TABLE OF CONTENTS

List of Figures .................................................................................................................... vi

List of Tables ................................................................................................................... viii

INTRODUCTION ...............................................................................................................1
   Goal Representations .................................................................................................. 2
   Counting: A Well-Learned Cascaded Task ............................................................... 3

FUNDAMENTAL ISSUES .................................................................................................7
   Storage Issues ............................................................................................................ 8
      Loading Working Memory .................................................................................. 10
      Over-Rehearsal .................................................................................................. 13
      Monitoring of Errors .......................................................................................... 15
   Piecemeal Construction of Intentions .............................................................. 16
      An Intention’s Point of No Return ..................................................................... 17
      Streamlining ....................................................................................................... 20

HYPOTHESES AND PREDICTIONS ........................................................................... 24

EXPERIMENT 1 .................................................................................................................26
   Method ....................................................................................................................... 28
   Results ...................................................................................................................... 29
   Discussion ............................................................................................................... 35

EXPERIMENT 2 .................................................................................................................36
   Method ....................................................................................................................... 40
   Results ...................................................................................................................... 42
   Discussion ............................................................................................................... 47

EXPERIMENT 3 .................................................................................................................50
   Method ....................................................................................................................... 52
   Results ...................................................................................................................... 54
   Discussion ............................................................................................................... 57

EXPERIMENT 4 .................................................................................................................58
Method .............................................................................................................. 60

Results ............................................................................................................. 62

Discussion ........................................................................................................ 68

GENERAL DISCUSSION ....................................................................................... 70

Explanation of Overcounting ............................................................................. 71

Overestimation Strategy Account ..................................................................... 71

Over-Rehearsal Account .................................................................................. 72

Distracters Account .......................................................................................... 73

The Nature of Goal Representations ............................................................... 75

Storage of Goal Representations .................................................................... 75

Linking goal representations .......................................................................... 80

Triggering of intentions ................................................................................... 82

Other Considerations ....................................................................................... 83

Individual differences of working memory capacity .................................... 83

The representation of numbers ...................................................................... 85

ACT-R .................................................................................................................. 86

Error detection ................................................................................................. 87

Limitations of studies ...................................................................................... 88

Future Directions ............................................................................................. 90

CONCLUSIONS .................................................................................................. 91

REFERENCES ................................................................................................... 96

APPENDIX A. ACT-R MODEL ........................................................................... 103

APPENDIX B. DISTRIBUTION OF ERRORS GREATER THAN FIVE ............. 118

APPENDIX C. TABLES OF ALL STATISTICAL TESTS .................................... 123

APPENDIX D. EXPERIMENT 2 RESULTS WITHOUT LOAD ERROR FILTER .... 131
LIST OF FIGURES

Figure 1. Representation of the mental processes of counting in approximate temporal synchrony with the events of counting (reproduced from Carlson & Cassenti, in press) .......................................................................................................................... 5

Figure 2. Proportion correct as a function of SOA condition ........................................ 30

Figure 3. Proportion of error trials with low and high confidence as a function of SOA condition ................................................................................................................................. 32

Figure 4. Proportion of correct trials with low and high confidence as a function of SOA condition ................................................................................................................................. 33

Figure 5. Coefficients of variation versus mean count response on double log coordinates. The lines correspond to the regression lines of the points for each SOA condition. The dotted line represents the 400 ms SOA condition and the full line represents the 3200 ms SOA condition. Of the two dashed lines, the smaller dashes represent 800 ms and the larger dashes represent 1600 ms ........................................... 34

Figure 6. Hypothesized sequence of events between response order conditions in a report total first condition. Goal representations to report the total are formed after the participant encodes the letter load in the blocked response order condition and are formed after the participant perceives the response cue in the mixed response order condition. A blocked response order condition requires storage of both count and report goal representations in working memory during counting, whereas only the count goal representation is stored in the mixed response order condition during counting .......... 38

Figure 7. Proportion correct as a function of SOA and Letter Set-Response Order condition .................................................................................................................................. 43

Figure 8. Proportion of error trials with low and high confidence as a function of SOA condition .................................................................................................................................. 45

Figure 9. Proportion of correct trials with low and high confidence as a function of condition .................................................................................................................................. 46

Figure 10. Letter-Typing response time as a function of SOA and letter set condition .................................................................................................................................. 47

Figure 11. Proportion correct as a function of distracter condition ........................................ 55

Figure 12. Proportion correct as a function of SOA and Press Instruction conditions ......... 62

Figure 13. Proportion of error trials with low and high confidence as a function of Press Instruction-SOA condition .................................................................................................................................. 65
Figure 14. Proportion of correct trials with low and high confidence as a function of Press Instruction SOA condition………………………………………………………………65

Figure 15. Median absolute error of presses from correct for each secondary task and SOA condition……………………………………………………………………………………67
LIST OF TABLES

Table 1. Frequency of undercount and overcount errors in Experiment 1………………31
Table 2. Frequency of undercount and overcount errors in Experiment 2………………44
Table 3. Frequency of undercount and overcount errors in Experiment 3………………56
Table 4. Frequency of undercount and overcount errors in Experiment 4………………63
INTRODUCTION

As the time to perform action decreases, so does performance on a task. This well-known finding is called the speed-accuracy tradeoff (Hale, 1969). If an individual increases the speed of performance, then there is a characteristic decline in probability of accurate performance. The focus of the current study is the effect of processing time that exceeds the amount of time needed to perform an activity.

There are several reasons to believe that more time does not always help performance. For instance, short-term memory must store task-relevant information in order to perform a task correctly. Forgetting is a problem for short-term memory and causes loss of information. Decay (see Peterson & Peterson, 1959) is a forgetting mechanism, which increases the chance that an object will be forgotten as more time passes. If decay results in loss of task-relevant information, then people will perform worse when given more time than needed to perform a mental operation.

Einstein, McDaniel, Williford, Pagan, and Dismukes (2003) found that forgetting is a substantial problem even over a period of several seconds. Their analysis indicated that people often forget their intentions to perform an activity even with short delays of only five seconds, resulting in decreases in performance. Given that all deliberate action begins with an intention to act, forgetting intentions can cause many problems (e.g., a pilot who forgets to lower landing gear when landing).

Clearly, there are limitations to the ubiquity of the speed-accuracy tradeoff. Forgetting is just one of the reasons why more time might hurt performance. I will follow-up on Einstein et al.’s (2003) suggestion by beginning my discussion of this problem with a summary of how goals and intentions impact cognitive processing. Next,
I will review the task used in the present study, event counting, and proceed to discussing possible reasons for how more time can negatively impact counting performance. Four experiments test these reasons (among other important issues) and will be followed by a discussion of the results.

*Goal Representations*

To achieve a desired state, an individual may need to accomplish a series of steps (e.g., counting all the members of a set) and each step begins with a goal representation. According to Carlson (1997, 2002), a goal representation is a schema or an organized set of elements that lists all necessary parameters required to complete a desired mental operation. The parameters that fit into the slots of the goal representation are the desired state (what the individual is acting to achieve), the operator (the type of action to be performed), and the operands (the objects which the operator acts on). For instance, an individual with a goal to solve an arithmetic problem would have to know that the desired outcome is to calculate a numeric result, the operator applied to achieve the outcome is adding, and the operands used by the operator are the displayed numbers.

Before the goal representation has all the necessary elements or parameters, it is a schema (i.e., a set of organized slots containing information important to a current task) with unfilled slots. A goal representation becomes an intention when all the parameters required to perform the goal are assimilated into the schema. An intention has the following general structure - \{I\} intend that \{outcome x be accomplished by {my} performing operation y on object(s) z\} (Carlson, 2002). Only the bracketed portion of the intention structure or goal representation will be considered here.
Intentions guide deliberate action (Carlson, 1997, 2002) and are essentially self-instructions for planned activity. I assume that goal representations are stored and assembled in working memory (see Altmann & Trafton, 2002).

Norman (1981) suggested that researchers can learn a lot about cognition by investigating errors and relating them to mental processes between intention and execution (as opposed to the errors in executing the operation). In terms of the present discussion and following the advice of Norman, I will relate errors to problems in the representation and storage of goal representations.

Norman (1981) suggested that an important next step in research on error was to investigate error experimentally rather than through naturalistic observation that had dominated the literature when Norman wrote his article. Since then, a majority of the experimental research on error between intention-formation and execution has focused on language production (see Postma, 2000 for a review) and not errors in other tasks (Byrne, 2003). Language production is one part of a large set of human skills and activities. New studies are necessary to explore errors in other skilled tasks to extend and develop theories of human error beyond the unique set of skills involved in language production (e.g., phonemic assembly processes, which construct the sound of words).

**Counting: A Well-Learned Cascaded Task**

A counting task allows an informative investigation of the structure and storage of intentions. Counting is a process with a series of intentions, one for each number in the counting sequence. Carlson and Cassenti (in press) detail a model in which the cascaded nature of these steps can account for errors in counting. Cascaded means that the
outcome from one step (the just-assigned total) serves as input to the next intention (to assign the next number up from the just-assigned total).

When an individual makes a counting error, it may take the form of an intention-outcome confusion. To illustrate, an individual may store the prospective (to be assigned) total in memory, intending to assign that number to the next item. When the next item appears, the individual may increment the total again by incorrectly labeling the intended, prospective total as the outcome of the last step and count from the prospective total (instead of with the prospective total). This mistake results in an overcount, an answer that is higher than the correct count. Similarly, an individual may undercount when an outcome total is mistaken for a prospective total and the individual fails to increment. Intention-outcome errors illustrate how the cascaded nature of counting is relevant to understanding the coordination of intentions.

Counting is commonly a skilled activity with a simple goal representation - that a new total be generated by assigning the next number to the new item. On any given step, the individual already knows the outcome should be a new total to be achieved by incrementing the current total. The only slot values not repeated on every step are the item and the prospective total. To fill in the prospective total slot of the goal representation, retrieval from long-term memory is required, which can occur soon after counting the previous item. Given sufficient time between counting events, all slots in the goal representation can be filled except the item slot.

The simple structure of counting allows an opportunity to model a minimal set of mental processes that an individual needs to count correctly. Figure 1 illustrates one such model by Carlson and Cassenti (in press), which uses Carlson’s (1997, 2002) conception
of goal representations. According to the model, plans are abstract representations of the goal of the next step (e.g., to assign the next number to the next event). A plan becomes an intention as its slots are filled with parameters (e.g., the prospective total fills the number slot). After the number and item slots are filled, the intention is complete and triggers a procedure. In counting, executing the procedure means consolidating the total in working memory and marking the event as counted. Once the new total is consolidated in working memory, the goal representation represents an outcome.

**Figure 1.** Representation of the mental processes of counting in approximate temporal synchrony with the events of counting (reproduced from Carlson & Cassenti, in press).

Several events take place in each counting step. First, the plan provides the schematic structure for the intention. Then, the prospective total fills the number slot, partially completing the goal representation. Attention focuses on the next event and
completes the intention by filling the item slot. Once the goal representation becomes an
intention, the procedure is enacted and an outcome representation is produced. The
outcome total is then active in working memory and provides a cue for retrieval of the
next prospective total, which fills the next prospective total slot.

provides a framework for implementing the above model. ACT-R is a production
system, a cognitive modelling system designed to represent a unified theory of cognition
(see Newell, 1990) or a theoretical framework that depicts all aspects of cognition under
one system of rules and constructs. Researchers implement the theory of ACT-R in
computer simulations of cognition to account for data from empirical studies. By
implementing ACT-R in this way, ACT-R develops stepwise and improves to account for
new phenomena in human cognition (e.g., Anderson & Lebiere, 1998). ACT-R is an
important system for the present research and will be referenced to illustrate
commonalities with the perspective taken here.

To coordinate goals, ACT-R uses a goal buffer to store goals, which are the
starting points for all action. The goal is a schematized structure that invokes a
production (like a procedure in the Carlson & Cassenti, in press, model). Productions
retrieve objects from declarative memory (i.e., long-term memory) and store results in a
retrieval buffer (like working memory). ACT-R/PM (PM stands for Perceptual-Motor,
Anderson, & Lebiere, 1998; Byrne, 2001) works in conjunction with ACT-R to perceive
objects in the environment through shifts in attention (analogous to the above model).
Given the close conceptual ties between ACT-R and the descriptive counting model,
Appendix 1 provides an account for how ACT-R may model counting along the assumptions of the descriptive model.

Another assumption of the counting model is that individuals represent numbers verbally. According to the model, individuals retrieve the prospective number from long-term memory. An alternative explanation is that individuals represent numbers with an analog representation (e.g., Gallistel & Gelman, 1992). An analog representation of number is like the markings on a ruler, the quantity of number is greater for the longer ruler markings. In other words, numbers may be represented visually, with a length marking quantity and quantity growing with each new count. Meck and Church (1983) suggested that animals represent number with an analog representation, but stopped short of suggesting that analog representation is the primary way that people represent number. Cordes, Gelman, Gallistel, and Whalen (2001) found that when people are free to use verbal representations of number they used them to count (this finding was based on a variability pattern characteristic of verbal representations of number). Nevertheless, Experiment 1 investigated the type of representation that people use to depict number.

FUNDAMENTAL ISSUES

Several issues are important for the analysis of why long delays may hurt performance. One important assumption is that working memory, the system that maintains and executes active information (for a review, see Miyake & Shah, 1999), stores intentions as it does any other task-relevant information (see Altmann & Trafton, 2002). This assumption is important because it implies that intentions are subject to the same types of degradation from storage as any other representations in working memory (i.e., decay and
interference). Errors made in well-learned tasks may be due to memory problems for either operands (e.g., the current count) or the intentions themselves.

Another probable source of error involves the structure of intentions. The goal representation contains multiple slots requiring specific content. In order to execute the to-be-performed operation correctly, intentions (as self-instructions) require the correct information. For instance, if an individual counts with the wrong number filling the prospective total slot, then the final answer will be incorrect, unless the mistake is corrected before the end of the sequence.

The storage and structure of goals are also properties of ACT-R. ACT-R goals are stored in the goal buffer where different parameter values of the goal representation are filled. Retrieval of parameters can come through ACT-R/PM’s vision module (e.g., the item operand) and through declarative memory retrieval (e.g., the prospective total). Storage of intentions in working memory and the structure of intentions are addressed in the next two sections.

**Storage Issues**

Norman (1981) discussed forgetting of goals as one of the major causes of error in well-learned tasks. Norman classified these errors as *loss of activation* errors. The classic example of this type of error is an individual who moves to a new room with a specific goal in mind and then ends up staring around the room wondering why he or she went there. Norman (1981) suggested that these types of errors are caused by interference and decay, common mechanisms of forgetting.

Altmann and Trafton (2002) adapted this perspective, suggesting that goals decay over time, making them more difficult to recall when needed. Altmann and Trafton
suggest that goal decay is functional in that it reduces interference with new goals. The
drawback of decay is that when the individual must retrieve the decayed goal after a
delay of up to a few seconds (Einstein McDaniel, Williford, Pagan, & Dismukes, 2003),
retrieval is more difficult.

If there are lengthy delays between items, an individual may prepare a new
intention to count long before the item appears. After filling the prospective total slot and
before the item appears, the only slot remaining unfilled is the item slot. The partially
complete goal representation is subject to decay while waiting for the next item and after
a long delay may be difficult to retrieve when needed.

This explanation accounts for a decrease in counting accuracy, but does not
predict the type of error made (overcount or undercount). A hypothesis proposed by
Buchner, Steffens, Irmen, and Wender (1998) implies that overcounting will be the most
common type of error. Buchner et al. suggest that participants generally overcount
because they forget the result of the last step and must reconstruct the count based on the
strongest memory trace of a total. Participants then use the strongest trace as the
minimum from which to reconstruct the count. Using this strategy, participants may
overestimate the total by attempting to compensate for any items counted since the
memory trace was first made. Perhaps decay makes retrieval of the goal representation
(containing the prospective total) more difficult and therefore individuals tend to
overcount when they re-estimate the total, especially with a long step time. This
hypothesis was tested in Experiment 2.

Buchner et al.’s (1998) approach contrasts somewhat with the one taken here.
Buchner et al. assumed that participants forget the current total from working memory.
In contrast, I assume that participants may forget either the prospective total or the entire goal representation (which contains the prospective total).

I also assume that the memory retrieval of the new total from long-term memory is a simple associative process that does not require a separate goal representation from the intention to count (see Carlson & Cassenti, in press, for a similar assumption). A mechanism in ACT-R captures this property. Links in ACT-R connect and spread activation from one declarative memory object to a related object automatically without a goal to retrieve. A counting model in ACT-R would set strong links and spreading activation from one number to the adjacent higher number.

*Loading Working Memory.* A test of the hypothesis that errors are caused by forgetting involves asking participants to count while holding a working memory load. A working memory load increases the chance of forgetting, forcing subjects to re-estimate the prospective total more often.

Here, I assume a hybrid model of working memory (see Nairne, 2002), which includes both activation-based forgetting (e.g., ACT-R) and interference. A hybrid model of working memory assumes a memory storage component that is limited in capacity by the total amount of activation available at any one time (e.g., the chance of forgetting is increased by having several highly-active items in memory) and assumes confusions between similar items in storage (i.e., interference). I also assume (like Altmann & Trafton, 2002) that goal representations are stored in working memory like any other task-relevant information. According to Altmann and Trafton (2002), old goal representations decay to prevent them from interfering with new goal representations.
This hypothesis suggests that when other goal representations persist in memory, they will interfere with a primary goal representation (such as counting).

Norman (1981) suggested that the interference from two active goals can cause forgetting of one of the goals (called unintentional activation errors). Heise, Gerjets, and Westermann (1997) demonstrated that a goal to perform a future task hurts performance of a well-learned task (such as counting). Interference from a secondary goal may be detrimental in two ways. First, any extra working memory load will make short-term remembering of any content more difficult (such as in ACT-R, Anderson & Lebiere, 1998). Second, because two goal representations are structurally similar, an intention to perform a secondary task may interfere with a primary goal more than any other type of memory load. The greater interference of one goal on another would be analogous to the greater interference of similar phonological loads (e.g., mat, that, and hat all sound alike) on one another (e.g., Baddeley, 1986). If the above analysis is correct, perhaps goal representations interfere with each other more because they are both stored with a similar structure. Experiments 2 and 4 explored this possibility.

Baddeley’s model (1986, also described in Baddeley & Hitch, 1974 and Baddeley & Logie, 1999) consists of three components: the central executive, the phonological loop, and the visuo-spatial sketchpad. The phonological loop stores verbal representations and the visuo-spatial sketchpad stores visual representations. The central executive stores small amounts of either type of information, but also engages in executive control functions. Baddeley assumed different components of working memory because two tasks theorized to use only the central executive, phonological loop, or visuo-spatial sketchpad can be performed at the same time without decreased
performance of either task, whereas performance of the same type of task decreases performance (e.g., Brooks, 1968, see Baddeley, 1986 for a review).

The present perspective assumes a unified working memory with different types of representations. Instead of separate components of working memory, I assume that all representations are stored in the same working memory store and that representations of the same format interfere (are confused) with one another. I propose that the strong interference between tasks that Baddeley assumes are performed in the central executive is actually interference between stored goal representations.

Carlson and Yaure (1990) demonstrated the importance of competing goal representations in a dual-task situation by investigating the effects of a secondary task on learning a primary task. While participants practiced logic rules, they were instructed to perform judgments (i.e., are two letters in different cases the same letter?), arithmetic problems, or remember words (short-term or long-term) between trials. Carlson and Yaure argued that the first two secondary tasks involved the formation of new intentions and therefore, required the participant to reconstruct operations for the logic rules before performing the primary task again. Reconstructing the operations resulted in better learning than remembering words, which merely required rehearsal. Carlson and Yaure demonstrated a benefit of secondary intentions in learning to solve logic problems, but also showed slower performance on the primary task during the acquisition phase of the experiment. Presumably, when goal representations compete for activation in working memory a similar disruptive effect occurs during counting.

A goal representation for a different task should compete with a goal representation to count, whereas a verbal working memory load should compete with the
prospective total, the verbal representation in counting. An experiment testing whether individuals forget the entire intention to count or just the prospective total requires two independent variables, a goal representation load either stored in working memory or long-term memory (storing in long-term memory does not interfere with working memory) and a verbal load in working memory or long-term memory. If performance declines because of a goal representation load, then perhaps people forget the goal representation to count and if performance declines with a verbal load, then perhaps people only forget the prospective total (e.g., Buchner et al., 1998). If performance does not decline with either type of load, then perhaps forgetting has little effect on counting. This reasoning was the foundation of Experiment 2.

Another consideration that may affect counting is the size of the individual’s working memory capacity (e.g., Daneman & Carpenter, 1980; Engle, Kane, & Tuholski, 1999). If some individuals can store more information in working memory than others can, then those individuals should be affected less by a working memory load. An aim of the present study was to investigate whether or not tests of working memory capacity (e.g., Daneman & Carpenter, 1980) could predict performance on a counting task with an intention load. Experiment 4 included a working memory span test to check on this possibility.

Over-Rehearsal. If memory decay affects the prospective total, then mental rehearsal would be an effective means of bolstering the strength of the weakened memory (e.g., Atkinson & Shiffrin, 1968). Mental rehearsal is a strategy of mentally repeating a representation with the aim of increasing the chance that it will be retrieved. The above discussion indicates that long delays between to-be-counted items will result in more
decay and thus more forgetting. If participants use rehearsal, then they may reverse the effects of decay during a long stimulus onset asynchrony (SOA or the time between the onset of one item and the onset of the next item) and increase the activation of the prospective total and goal representation.

Carlson and Cassenti (in press) suggest that intention-outcome confusions are a possible cause of overcounting in their studies. As discussed above, when the prospective total is confused for an outcome, the participant may incorrectly label the prospective total as the current total and thus commit an intention-outcome confusion. This type of intention-outcome confusion results in an overcount.

One method of creating an intention-outcome confusion is to raise the activation strength of the prospective total so high that it is confused as the current total. If participants rehearse the prospective total too many times (to prevent decay of the total or goal representation), then they may commit intention-outcome confusions. Intention-outcome confusions would therefore be more likely in the long SOA condition, in which rehearsal is more necessary to prevent the adverse effects of decay than in a short SOA condition.

Nairne (2002) provides an alternative point of view supporting this prediction of over-rehearsal. Nairne suggests that rehearsal does not raise the activation strength of a representation, but instead creates multiple copies of the representation and increases the chance that the representation will be recalled though one of its duplications. If rehearsal creates multiple copies of the prospective total and one of those copies is mistaken for the outcome representation (i.e., the current number), then an individual will be more likely to commit an intention-outcome confusion with greater levels of rehearsal.
The over-rehearsal account and the forgetting-overestimating theory of Buchner et al. (1998) both predict overcounting during a long SOA. An experiment to distinguish between these two hypotheses requires a working memory load. If memory stores are devoted to more representations, then the chance to rehearse decreases and therefore the chance of forgetting increases (see Rundus, 1971). If working memory loads cause more overcount errors, then the forgetting-overestimate theory is the more probable cause of error. If working memory loads do not lead to more overcount errors, then the over-rehearsal explanation of overcounting is the more probable cause of error. The results of Experiment 2 were interpreted with this reasoning.

*Monitoring of Errors.* If a counting error occurs because of forgetting the prospective total, then the individual has no number to draw from working memory when trying to assign a number to a new event. Not having a number to draw from memory is a detectable error that necessitates re-estimation of the prospective total. Therefore, an error caused by forgetting likely results in awareness of the error and successful error monitoring, but not necessarily a successful correction. Forgetting is more likely during long intervals, so forgetting the total should occur more often when there is a long wait time between retrieving the prospective total and the appearance of the next item (when the intention can trigger and use the prospective total). Confusing an intention for an outcome should also be a noticeable disruption to counting because typically individuals should readily distinguish current and prospective number. Therefore, over-rehearsal should result in more detected errors. In contrast, individuals should not detect an error when one and only one prospective number is selected from memory for every new item that appears.
Monitoring task performance is an important construct that was examined in all of the present empirical studies. By rating their confidence in reported totals, participants should demonstrate awareness of the errors caused by forgetting or over-rehearsal, by responding with low confidence. I predict that participants will be more successful in rating their own errors with low confidence when wait time for new events is longer and forgetting or over-rehearsal is more likely to occur. This prediction was tested in Experiment 2 and 4.

*Piecemeal Construction of Intentions*

Other properties of goal representation are important to the question of how long delays may hurt performance. The triggering condition of an intention may be an inherent property of intentions or an external mechanism acting on the intention (e.g., a central executive, see Baddeley, 1986). I assume that the triggering condition is an inherent property of intentions. Not only does an external mechanism introduce other important questions (e.g., how to resolve the homunculus problem, see Carlson, 1997), but also if the intention waits in storage for an external mechanism, then the chance of forgetting increases. If intentions suffer from the same memory problems as other items in memory, then intentions should trigger and execute as soon as they are completed (see section on the point of no return, below). The longer a completed intention must stay in memory, the greater the chance of forgetting (e.g., Altmann & Trafton, 2002). Therefore, important issues are when and how intentions are completed - issues involving the structure of intentions.

Carlson (2002) described the structure of intentions as schemas with slots for specifying the key components of a mental operation. Specifically, the slots of the goal
representation are the desired outcome, the operator (i.e., type of action), and the operands (i.e., objects used by the operator). Together, the slots specify an intention or self-instruction to act, which may be part of a series of actions used to achieve a broader goal (e.g., counting one event in order to count all the events).

An intention is constructed as the focus of attention shifts and fills each slot in the schematic goal representation. This piecemeal construction of an intention is an important part of the dynamic role of intentions in cognition. If individuals were to fill in all slots simultaneously, then the coordination involved would tax cognitive resources and cause more errors than piecemeal construction. I assume that goal representations are filled in piecemeal - when one parameter becomes available, it immediately fills the slot in the goal representation where it is stored until the remaining pieces of the goal representation are filled.

The piecemeal construction assumption suggests an important question for study - if goal representations are filled piecemeal, then when does an intention trigger to execute the specified operation? One answer is that the intention triggers a procedure when all the slots in the goal representation are filled. If goal representations are subject to decay and interference, then execution at time of completion seems the most appropriate triggering condition to avoid the detrimental effects of forgetting. Delaying the intention trigger after completing the intention may hurt performance because of greater decay with longer storage time. Triggering the intention before completion would execute an underspecified intention and lead to a failed operation.

*An Intention’s Point of No Return.* To anticipate the results, Experiment 1 showed overcounting at long rhythmic intervals between to-be-counted events.
Overcounting during rhythmic counting contrasts with previous results showing undercounting of rhythmic events at faster paces (see Carlson & Cassenti, in press). The following account provides an interpretation of this result, emphasizing the structure and storage of goal representations.

One explanation for overcounting starts with the assumption that goals have a point of no return, after which they trigger and execute the specified operation. Consider the goal representation portion of an intention, that outcome x be accomplished by performing operation y on object(s) z. In the goal representation, there are three types of slots that need to be filled to complete the parameters of the goal: the desired outcome, the operator, and the objects or operands. Once each slot is filled, the intention may trigger and complete the operation. This process is analogous to a process in ACT-R, in which procedures are executed once all the important parameters of the triggering condition are filled. In order to avoid error, individuals must fill these slots with correct information.

This analysis can be elaborated based on research on stop signal tasks (for a review, see Logan & Cowan, 1984). In a stop signal task, participants are instructed to respond to presented stimuli but to withhold the response if a stop signal sounds (e.g., a beep). Logan and Cowan review evidence showing that subjects have a more difficult time stopping the response when the stop signal sounds just before the response. Their analysis implies a race between the mental processes involved in the primary response and the stopping response. They suggest that once the mental processes of the primary response pass a point of no return, the primary response executes ballistically (i.e., without the possibility of stopping). Logan and Cowan described the existence of
ballistic processing, but did not describe the locus of the point of no return. Perhaps, the point of no return for intentions to trigger and execute (including the intentions of participants in the Logan and Cowan study) is when the intention is complete.

Consider counting, a task with a repeated goal representation that a new total be accomplished by assigning the next number to the next item. Before the appearance of the next item, the desired outcome (a new total), one operand (the next total), and the operator (to assign a number) are already available to complete a new goal representation. The only missing operand is the to-be-counted item. Between items, participants confront a number of internal (e.g., self-generated thoughts unrelated to the current task, see Kuhl & Koch, 1984) and external (e.g., background noise) distractions, one of which might be mistaken for the to-be-counted item. If a distracter fills the item slot, then the distracter will complete the goal representation and trigger the intention.

Reason (1979) attributes many errors in well-learned tasks to misdirection of focal attention. According to Reason, information in focal attention can trigger motor programs inappropriate to the current goal. An assumption of the present research is that focal attention can also disrupt processing of the current intention by triggering the intention prematurely. For example, a distracter stimulus might be mistaken for a counting target and trigger the next count before the appearance of the next target.

Carlson and Cassenti (in press) provided some evidence for the hypothesis that stimuli other than targets can trigger intentions to count. In one study, a probe number appeared after the final item with the rhythmic interval of time when another item would have appeared. The participants were instructed to decide whether the probe number was the current total. Participants tended to respond that the probe number was correct when
it was actually one higher than the current total. This result suggested that the probe number triggered the intention to count, resulting in an overcount. In the present study, Experiments 3 and 4 checked this possibility by presenting distracter stimuli during delays between items. The expectation was that distracters would cause more overcount errors.

Most of the stop signal studies involve the study of motor processes (e.g., Osman, Kornblum, & Meyer, 1986) not simply mental processes. Two notable exceptions are Logan (1983) and Zbrodoff and Logan (1986), who studied word recognition and arithmetic, respectively. The procedures of these two studies involved presenting a problem with or without a stop signal and later testing the participants’ memory for stimuli of stop and no stop trials. The results showed equivalent memory for both late signal stop and no stop signal trials, suggesting that, even when stopped, participants mentally processed the problems enough to form strong memories of them.

A counting task provides an opportunity for further examining the ballistic nature of mental processes. Participants could inhibit responses after stop signals in the above studies but showed ballistic mental processing through strong memory of stopped problems. If distracters are presented during counting, ballistic processing can be tested more directly. If more distracters show more overcounting, then sometimes participants confuse distracters for targets and therefore improperly trigger and execute intentions.

Streamlining. An important consideration for the hypothesis that distracters sometimes erroneously complete goal representations is streamlining. As individuals gain skill, they also learn to ignore unnecessary perceptual information (e.g., Haider & Frensch, 1996, 1999). I will refer to the process of focusing on only task-relevant
information as streamlining. Consider an assembly line worker who must insert batteries into a battery-operated device. At first, this worker may reference the picture on the inside of the battery case to fit the batteries in the correct direction. After many trials, the worker no longer needs to reference the picture because the batteries always fit in the same direction. Ignoring the battery pictures is an example of streamlining the assembly process.

Haider and Frensch (1996) discovered that individuals learned to respond to complex problems faster with practice. They attributed the speed-up to the participant’s ability to ignore unimportant information. In the task, information that did not affect the answer was presented in the first part of the study, so participants learned to ignore it. Haider and Frensch demonstrated a cost to streamlining information in the second part of the experiment. After the participants learned to ignore the information, practiced participants who were not receiving accuracy feedback continued to disregard the information even when ignoring it caused them to answer incorrectly. Participants who received accuracy feedback initially ignored the now-irrelevant information, but learned to attend to the information after a number of trials. To further test the information reduction hypothesis, Haider and Frensch (1999) tracked participants’ eyes while they performed the same task. Haider and Frensch (1999) replicated both findings described above. They also discovered that participants focused their eyes less on the unimportant information after a number of trials and focused primarily on the parts of the problem that changed.

Correct counting requires one-to-one correspondence (Gallistel & Gelman, 1992) of items to number totals. Errors occur when numbers and items do not match up one to
one (e.g., one item gets two numbers or no number). To achieve one-to-one correspondence, the individual must perceptually distinguish the current item from the previous item and from the next item. Distinguishing items helps the individual count, although distinguishing items will not necessarily lead to the correct count (i.e., the wrong numbers may be assigned).

In event counting, each item appears one at a time. The time between items (a blank screen) indicates that two items are distinctive from one another and is the minimum amount of information needed to distinguish one event from another. Item identity (e.g., distinctive perceptual features or category membership) could also distinguish one item from the next. An individual who counts distinctive items may either streamline encoding by ignoring identity information or keep track of this information and use it to distinguish adjacent items.

Streamlining in counting has both benefits and costs. To achieve one-to-one correspondence, an individual must attend to both the number sequence and the item sequence. If participants ignore item identity, then the amount of information is reduced, but distinguishing between the current item and the previous item is more difficult (if two items have different identities then they are distinct from one another). When counting identical items that are displayed all at once, an individual can distinguish items based on their location. In event counting, identical items that all appear in the same location can only be distinguished by when they appear.

A cost of distinguishing items based solely on when items appear is that the individual cannot recover ignored information, once the display changes. If perception of the blank screen separating events fails (e.g., attention wanders), distinct identities for
consecutive items may help the individual keep track of the sequence of items (e.g., the current item has a different identity from the previous item). Underwood (1982) suggests that feedback is an important part of any activity. Feedback provides information about the success of a particular operation. Distinctive item identities may serve as a basis for self-generated feedback. If the identity of the previous item was different than the current item then identities provide feedback that the previous intention was executed and a new intention is needed.

This self-generated feedback may be important in ensuring that only correct items (i.e., targets) and not incorrect items (i.e., distracters) fill the item slot in the intention to count. In order to prevent a distracter from filling the item slot, individuals must monitor the formation of the goal representation. If individuals count identical items (e.g., asterisks), they are likely to ignore the identity of the items because the identities never change. Without identity information, individuals may count whenever they see a contrast change in the blank background (of the screen). If the identities of identical items are ignored, it may be easier to include a distracter as the item operand when counting identical items versus distinctive items.

Perhaps when participants apply a different identity to each successive item they focus more attention on which item is presented and are more careful about which item fills the item slot. The design of Experiment 3 consisted of distracters presented during the presentation of distinctive items and distracters presented during identical items. The prediction was that identical items would cause more overcounting than distinctive items because of a greater susceptibility to distracters with increased streamlining.
HYPOTHESES AND PREDICTIONS

The background information outlined above provides an account of how and why too much processing time can disrupt performance in cognitive tasks. The premise is that goal representations, which are necessary for deliberate action, have properties that make performance decline when processing time is longer than optimal. First, goal representations are subject to the same types of memory problems as any other representation (e.g., Altmann & Trafton, 2002). These memory problems may cause forgetting of goal representations or cause a strategy of over-rehearsal, which also leads to errors. Second, goal representations may have a point of no return, which causes intentions to trigger prematurely when presented with incorrect information. A review of these hypotheses is presented below.

The forgetting hypothesis suggests that because goal representations are subject to decay, the longer a goal representation is stored in working memory, the greater the chance that the goal representation will be forgotten. In terms of counting, this suggests that the goal representation may need to be revived during a long step time (SOA) along with the prospective total (contained within the decayed goal representation). According to Buchner et al.’s (1998) overestimation strategy, when individuals forget the current total they must re-estimate it and will choose a number that is higher than the one they first retrieved. Overestimation is a strategy invoked to compensate for any items that may have appeared since the memory of the number was formed.

A second hypothesis suggests that goal representations are most susceptible to interference from other goal representations because of their similar structures. If a verbal working memory load interferes with the verbal representation of the prospective
total, then a goal representation load interferes with the primary task goal representation. More forgetting should lead to more re-estimation of the prospective total and therefore more overcounting errors.

Over-rehearsal is another possible mechanism making counting with a long SOA more difficult. The over-rehearsal theory suggests that in order to overcome the detrimental effects of decay, which is greatest in a long SOA condition, individuals will rehearse the prospective total and increase its activation strength. If individuals boost the activation of the prospective total too high, then the prospective total can be confused with the current total, which may lead to an intention-outcome confusion. In other words, participants may overcount because they confuse an element of the intention, the prospective total, for an already completed count or outcome, mistaking the prospective total as the current total.

A fourth hypothesis suggests that incorrect information, which fills the last remaining slot in a goal representation can cause premature triggering of the intention, assuming that intentions trigger once they are completed. I assume that intentions have slots specifying the parameters of the intention and that these slots are filled in piecemeal (as opposed to all-at-once). Intentions should trigger at a consistent point in time to coordinate mental activity. The best candidate time for completing an intention is when the intention has been fully developed. Intentions cannot trigger before completion because the intended operation would be under-specified and if they trigger too long after completion, intentions are at risk for increased decay and interference. Therefore, the time at which an intention is completed is the intention’s point of no return – the point at
which the operation specified by the intention is triggered and runs to completion
ballistically.

If this hypothesis is correct, then incorrect information that completes the
intention will erroneously trigger the intention. When individuals count events with a
long SOA, both the desired outcome (i.e., a new total) and the operator (i.e., to assign a
new number) are available from the start. The prospective total operand can be retrieved
from long-term memory and filled in soon afterward. The only remaining slot to
complete the goal representation is the item slot. A distracter that fills in the remaining
slot will trigger the intention and erroneously complete the count. Presenting distracters
during counting should increase the amount of overcounting, if participants count some
of the distracters and all of the targets.

These hypotheses and predictions suggest several experiments to focus on the
nature of errors during a well-learned task with a long step time. The following describes
four experiments designed to investigate the important issues summarized above.

EXPERIMENT 1
The first experiment was designed to investigate the effect of a long event SOA in a
counting task. The hypothesis that long SOAs increase overcounting errors can be
explained by the forgetting hypothesis that intentions are stored in working memory and
experience the same type of decay and interference as other items in working memory.
Forgetting the intention should induce more overcounting errors (see Buchner et al,
1998). Because items that stay in working memory longer have a greater chance of being
forgotten, a long waiting period between counting events increases the chance of
forgetting and overcounting.
Alternatively, participants may engage in a rehearsal process to boost the strength of the prospective total to overcome the effect of decay. This strategy would reduce forgetting, but it may also create more intention-outcome confusions by raising the activation of the prospective total too high. Intention-outcome confusions lead to more overcount errors.

Another explanation for errors in counting suggests that distracters might fill the item operand slot in the goal representation to count. Assuming a constant rate of distracters in the environment (e.g., a flicker from the room light), the longer the event SOA the greater the chance that a distracter will appear and complete the intention. Together, the forgetting, over-rehearsal, and distracters hypotheses suggest that the longer a participant waits for the next event, the greater the chance of making an overcount error.

In Experiment 1, participants counted asterisks that appeared with rhythmic SOAs (i.e., time between the onset of one item and the next item) of 400, 800, 1600, and 3200 milliseconds (ms). Previous research (Carlson & Cassenti, in press) showed that when participants counted events appearing at rhythms up to 550 ms, they typically showed undercounting. The expectation was that the longest SOA condition would show more overcounting errors than the short SOA conditions because of the three mechanisms outlined above: increased forgetting from greater decay, more intention-outcome confusions from over-rehearsal, and more incorrect triggering of intentions from more frequent distracters.

Another goal of Experiment 1 was to investigate whether participants use verbal or analogical representations of number. Cordes et al. (2001) demonstrate that the
relationship between the log of the ratio of the standard deviation of responses to trials over the mean of the total number of items presented (known as the coefficient of variation) and the log of the mean count response reveals the type of number representation participants use to count. If a regression line for the log of the coefficient of variation and the log of the mean count response has a zero slope then, participants likely counted using an analogical representation of number and if the slope is -0.5 or lower than participants likely used a verbal representation of number.

Method

Participants. Twenty-four students in an introductory psychology course at the Pennsylvania State University participated in exchange for a small amount of course credit.

Materials. All experiments were run on a Windows-based personal computer. The experiment was controlled by E-Prime, version 1.0 and participants used the attached keyboard to enter their responses.

Design. The lone independent variable was the SOA of the counting events. Five trials each of 400, 800, 1600, or 3200 ms SOA were chosen randomly without replacement for a total of 20 trials for every participant.

Procedure. Participants were instructed to count the number of asterisks that appeared on the screen and to enter their response at the end of the sequence. After entering their count response, the participants were instructed to enter the confidence that their answer was correct on a scale of 1 to 5. Participants were also instructed to keep both hands flat against the table in front of them while counting to avoid external strategies that aid counting (e.g., pointing at the number keys on the keyboard to keep
track of the one’s digit of the count total). An experimenter remained in the room to
ensure that this constraint was satisfied. All responses were typed into Windows
response boxes requesting the correct count and confidence (on a scale of 1 to 5). Non-
numeric responses or responses out of the range of 1-5 for confidence triggered requests
for re-typing.

Before the start of the trial, a random number between 30 and 80 was chosen as
the number of steps in the counting sequence. Participants began the trial by pressing the
space bar at a prompt screen. After the space press, the screen was blank for a time equal
to the SOA condition minus 200 ms. After the blank interval, an 18 point Courier New
font white asterisk appeared in the middle of the black background for 200 ms. After the
asterisk screen, the plain black screen appeared for the same length of time as the first
blank screen. The first response box appeared immediately after the 200 ms of the last
asterisk. Accuracy feedback along with the correct answer followed the last confidence
response and a wait screen of 10 seconds followed the feedback screen, to allow the
participants to rest their eyes. At the end of the wait, the screen flashed and the
participant pressed the space bar to begin the next trial.

After the last trial, a thank you screen appeared. Each participant was then
debriefed on the purpose of the experiment.

Results

All responses of a magnitude greater than five from the correct response were
excluded from all analyses. As in Carlson and Cassenti (in press), these responses were
considered as either typing errors or misremembering the ten’s digit. Both of these types
of errors are outside the scope of the theoretical claims made here. Errors of greater than
five accounted for 8.1% of all data in this study. The following accounts for significant findings of Experiment 1 and Appendix 2 contains all statistical test results for both significant and non-significant findings in Experiment 1 and all subsequent experiments.

**Accuracy.** Only responses that were the same as the correct total were considered accurate. All other responses were considered as errors. Participants made correct responses on 73.9% of all trials. Figure 2 shows the distribution of proportion correct as a function of SOA condition.

![Figure 2. Proportion correct as a function of SOA condition.](image)

A one-way analysis of variance (ANOVA) on the four levels of SOA (400, 800, 1600, and 3200 ms) demonstrated a reliable difference in accuracy between SOA conditions, $F(3,69) = 11.08, MSE = 0.37, p < .001$. The linear term of accuracy and SOA was not significant, $F(94) = 1.35, RMS_E = 0.014, p < 0.25$, but the quadratic term was significant, $F(93) = 5.77, RMS_E = 0.110, p = 0.004$. Planned t-test comparisons showed the 3200 ms SOA condition (0.739) was more accurate than the 400 ms condition (0.555), $t(23) = 2.45, SE_{M} = 0.075, p = 0.011$. The 3200 ms SOA condition was also less
accurate than the 800 ms (0.835), $t(23) = 2.03, SE_M = 0.047, p = 0.027$ and 1600 ms (0.827), $t(23) = 2.04, SE_M = 0.044, p = 0.029$ conditions (one-tailed tests).

*Signed Error.* Signed error was calculated by subtracting the correct answer from the participant’s response. If the participant entered a count that was too low, the response was an undercount. If the participant entered a count that was too high, the response was an overcount. A chi-square test was conducted on the frequencies of undercount and overcount errors. Because of the small pool of errors across participants, a chi-square test was chosen over an ANOVA to test the ratio of overcounts to undercounts within each SOA condition. Table 1 represents the frequencies of undercounts and overcounts in each SOA condition and the marginal totals.

Table 1.
Frequency of undercount and overcount errors in Experiment 1.

<table>
<thead>
<tr>
<th>SOA</th>
<th>Undercount</th>
<th>Overcount</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>27</td>
<td>19</td>
<td>46</td>
</tr>
<tr>
<td>800</td>
<td>10</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>1600</td>
<td>14</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>3200</td>
<td>9</td>
<td>18</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>49</td>
<td>109</td>
</tr>
</tbody>
</table>

*Note.* Only includes errors of five or less in magnitude.

All undercount and overcount errors were added across SOA conditions to determine the overall likelihood of each type of error. Overall, there was a 55.0 percent chance of an undercount and a 45.0 percent chance of an overcount. Each SOA condition was tested individually to check if the frequency of undercounts and overcounts differed from this likelihood proportion. Only the 3200 ms SOA condition differed from this proportion, $\chi^2 (1) = 5.12, p < .05$ demonstrating a greater number of overcount errors and a lower number of undercount errors than predicted.
**Error Monitoring.** Confidence ratings on a scale of 1 to 5 were divided into low confidence (ratings of 1 to 3) and high confidence (ratings of 4 and 5) for each SOA condition. A proportion of low and high confidence was calculated for both error and correct trials. Figure 3 displays this proportion for error trials. Figure 4 presents this proportion for correct trials.

![Figure 3](image_url)

**Figure 3.** Proportion of error trials with low and high confidence as a function of SOA condition.
Figure 4. Proportion of correct trials with low and high confidence as a function of SOA condition.

Figure 3 illustrates that when errors were made in the 3200 ms SOA condition the participants showed a substantially greater chance of recognizing that the error was made than the 800 and 1600 ms SOA conditions, by rating the confidence in the count response as low. Figure 4 illustrates that when errors were not made, all SOA conditions, except 400 SOA were rated with equally high confidence.

**Number representation.** Figure 5 presents data points of the log of the coefficient of variation (i.e., the ratio of standard deviation of count responses to the mean number of items presented) and the log of the mean count response across trials for each subject. Figure 5 also presents the regression lines representing the relationship between these two variables.
Figure 5. Coefficients of variation versus mean count response on double log coordinates. The lines correspond to the regression lines of the points for each SOA condition. The dotted line represents the 400 ms SOA condition and the full line represents the 3200 ms SOA condition. Of the two dashed lines, the smaller dashes represent 800 ms and the larger dashes represent 1600 ms.

The following are the regression line parameters for each SOA condition: 

- For 400 ms: \(y = 0.001x - 0.577\), \(r = 0.001\)
- For 800 ms: \(y = -0.839x + 0.862\), \(r = 0.300\)
- For 1600 ms: \(y = -1.807x + 2.517\), \(r = 0.480\)
- For 3200 ms: \(y = -1.102x + 1.272\), \(r = 0.408\)

The slopes of the 800, 1600, and 3200 ms SOA conditions were all well below zero with high correlations, indicating reliable differences from a zero slope. The 400 ms SOA condition had a slope of almost zero, however the correlation indicates low reliability for the conclusion that the points fit along a slope of zero.
Discussion.

These results are favorable to the predictions that counting performance declines at long SOAs and that those errors favor overcounting. First, accuracy was better for the intermediate SOA conditions of 800 and 1600 ms than for the long SOA condition of 3200 ms. Along with the significant quadratic term of the accuracy curve and SOA, these results support the claim that counting performance increases as SOA increases up to a point between 1600 and 3200 ms, after which it decreases with longer SOAs.

Second, the finding that overcounting in the long SOA condition was more frequent than predicted supports the hypothesis that long SOAs will result in more overcounting than shorter SOAs. The hypothesis that the prospective total is forgotten from working memory (via forgetting the intention) may explain the overcounting effect because the chance of forgetting is greater over longer SOAs (triggering re-estimation – a strategy that results in overcounting according to Buchner et al., 1998). Over-rehearsal increasing intention-outcome confusions is another potential explanation for overcounting. The hypothesis that distracters fill the item slot may also explain the overcounting effect because the chance of a distracter increases at longer SOAs. Follow-up experiments are necessary to discriminate the impact of these potential causes.

Additionally, participants demonstrated substantially better error monitoring in the long SOA condition than the other SOA conditions. All main hypotheses predict improved error monitoring at a long SOA. Forgetting, intention-outcome confusions, and distractions are disruptive, noticeable, and are more likely at a long SOA. Improved monitoring shows that either or both of these hypotheses may be true because a highly noticeable error is more likely to prompt lower confidence.
The final test investigated the type of representation participants used to represent number. The results indicate similar variability signatures as found in verbal counting of the Cordes et al. (2001) findings for all SOA condition above 400 ms. The finding for the 400 ms SOA condition was inconclusive. All remaining experiments used only 800 and 3200 ms SOA conditions and assumed that participants used a verbal format to represent numbers.

EXPERIMENT 2

Experiment 1 demonstrated more overcounting and lower accuracy for the longest SOA condition. However, no one cause for overcounting emerged from the results of Experiment 1. The three hypotheses posed above that overcounting results from forgetting, over-rehearsal, or distracters incorrectly triggering intentions are all possible causes. In order to begin discriminating these three causes, Experiment 2 was designed to focus on the forgetting and over-rehearsal hypotheses.

To increase the chance of forgetting and decrease the chance of rehearsal, all participants engaged in a secondary memory task while counting. All participants were asked to remember a set of letters and to type the set after counting. Half of the participants remembered the same set of letters on every trial and the other half remembered a different set of letters on every trial. Participants who remembered a different letter set had to store the set in working memory while counting and would therefore be more susceptible to forgetting other verbal representations, in this case the prospective total. Participants who remembered the same letter set on every trial, merely stored the set in long-term memory, requiring less working memory capacity.
A second manipulation involved a goal representation load. In Experiment 2, participants typed two responses at the end of each trial, the letter set and the count. Half of the participants were informed as to which response would come first and that response order would stay the same over the next several trials. I will refer to this condition as the blocked response order condition. The other half of the participants were not informed as to which response would come first and the sequence of response order trials was randomly mixed. I will refer to this condition as the mixed response order condition. Trials in which participants typed the letters first may have resulted in more forgetting of the total. Because counting performance was the focus of Experiment 2, only count first trials were considered in this experiment.

All participants were asked to report both the memory load and count as quickly as possible (when prompted) and were given feedback on how quickly and accurately they responded. Because the sequence of events ends unpredictably, the participants with a blocked response order were inclined to store a second goal representation (i.e., to type the letters) while storing the goal representation to count because they would be ready to type the response as soon as the first response request was made. The mixed response order participants were less likely to keep the letter-typing goal representation active while counting because they were not informed of which response would come first. Even if mixed response order participants were to store two other goal representations, to type the total and to type the letters, they would still need to check the request cue before typing. Therefore, the prediction was that the blocked response order would cause faster responses than the mixed response order. However, accuracy for the blocked response order group should decrease relative to mixed response order because participants in the
blocked response order had a greater working memory load. Figure 6 depicts the sequence of events resulting in a goal representation load during counting with a blocked response order and no goal representation load with a mixed response order. The prediction was that the report goal representation load would decrease counting performance for the blocked condition.

Figure 6. Hypothesized sequence of events between response order conditions in a report total first condition. Goal representations to report the total are formed after the participant encodes the letter load in the blocked response order condition and are formed after the participant perceives the response cue in the mixed response order condition. A blocked response order condition requires storage of both count and report goal representations in working memory during counting, whereas only the count goal representation is stored in the mixed response order condition during counting.

This procedure tested two hypotheses. First, it tested whether loading working memory with a verbal load or a goal representation load increases the chance of forgetting and if increased forgetting raises the number of overcounts made. Loading
working memory with a verbal load or a goal representation load should increase forgetting and therefore lower accuracy. According to the theory by Buchner et al. (1998), if there is an increase in forgetting there should be a corresponding increase in the number of overcounts.

Second, it tested whether a working memory load would decrease rehearsal of the prospective total and therefore decrease the number or intention-outcome confusions. If participants make fewer intention-outcome confusions, then the number of overcount errors should decrease even when events appear with a long SOA. If participants rehearsed the working memory load then they would have rehearsed the prospective total less often than if there was no working memory load.

All participants had trials of 800 and 3200 ms SOA. In Experiment 2, the activation of the working memory load (i.e., rehearsing the letter set and the goal of typing the letter load) should be greater for the long SOA condition because participants have more time to activate the load while counting. If activating the secondary task increases interference with the counting task, then participants should make more errors with a long SOA.

Overall, counting accuracy should show an interaction between memory load and interference. As the working memory load increases from neither letter set nor goal representation, to one of the two, to both, count accuracy should decrease. The amount of disruption should be highest in the long SOA condition because the working memory loads should cause more forgetting in the condition that also suffers from the most decay and because the load will hurt the ability to rehearse (which causes more overcounts, but reduces the amount of forgetting).
Method

Participants. Sixty-five students in an introductory psychology course at The Pennsylvania State University participated to fulfill a small amount of course requirement. One participant was removed from all analyses for opting to exit in the middle of the experiment.

Design. There were three independent variables. First, SOA of counting events varied within-subjects. Ten trials for each participant were an 800 ms SOA and the other ten trials were a 3200 ms SOA, for a total of 20 trials for each participant. The other two independent variables were manipulated between-subjects. Thirty-two participants remembered the same letters set throughout all the trials. The other thirty-two participants remembered a different letter set on every trial. Half of the participants in each of the letter set groups responded to blocked response order trials and the other half of participants in each group responded to mixed response order trials.

Procedure. The display showed a list of seven letters to begin each trial. The list appeared in Courier New 18 point white font in the center of the screen, listed vertically. Participants were instructed to press the space bar to remove the list and begin the counting task once they had memorized the set.

The counting task was the same as Experiment 1, except instead of four SOA conditions, there were only two SOA conditions – 800 and 3200 ms.

After the final asterisk, a screen display appeared requesting the count total or the letter load. The request appeared in 26 point Courier New font above where the participant entered the response. The word “Letters” appeared if requesting the letter load or “Number” if requesting the count total. If the trial was a blocked response order
the first response was always the one that the participant had been informed would be requested first. Before the first trial and after every five trials participants in the blocked condition were informed of a new response order. If the participants typed one type of response first, the next set of five trials was always the other type of response. Half of all participants in the blocked condition began with the letter response first and the other half began with the total response first. In the mixed response order condition, the participants saw no screen informing them of the first response. The order of first response trials was random with the condition that half of all trials made the letter request first and the other half were total request first.

The screens requesting the memory load and count response were changed from Experiment 1. At the bottom and to the left of each response screen a vertical line (i.e., “|”) appeared. The vertical line acted as a cursor from which the participant could type the response. After each key press, the response was displayed in 18 point Courier New font and pushed the cursor to the right. When reporting the memory load a space appeared between each letter and in the count report both numbers appeared side by side. Once the participant typed two numbers or seven letters, the cursor disappeared from the screen. Participants were instructed to hit the left arrow key to delete a typed character and to press enter when finished responding. Only displayed responses entered accuracy calculations (i.e., the participant could type beyond the allotted character limit, but only the first two digits or seven letters were considered).

After entering both responses, the feedback screen displayed the accuracy and response time of both the count and the memory load responses. The feedback screen also displayed the average time for each response throughout the trials. A ten second
wait followed the feedback screen to allow participants to rest their eyes. After the last trial, a thank you screen appeared and each participant was debriefed on the purpose of the experiment.

**Results**

All responses in which the letter set was not 100% correct were eliminated from all further analyses to limit the data set to only trials in which participants devoted enough working memory resources to the verbal load while counting to report the letters with complete accuracy. Error responses accounted for 19.2% of the data. Appendix 4 contains all results before this filter (with the absolute value of error filter below) and illustrates that the filter did not have a profound effect on the interpretation of the data.

All responses of a magnitude greater than five from the correct response were excluded from all analyses. As in Experiment 1, these responses were considered as either typing errors or misremembering the ten’s digit. Errors of greater than five accounted for 3.5% of the remaining data in this study.

**Accuracy.** Only responses that were the same as the correct total were considered accurate. All other responses were considered errors. Participants made correct responses on 83.0% of all trials. Figure 7 shows the distribution of proportion correct by SOA, letter set, and response order conditions.
A mixed analysis of variance (ANOVA) on the within-subjects levels of SOA (800 and 3200 ms) and the between-subjects levels of letter set (different and same) and response order (blocked and mixed) demonstrated a reliable difference in counting accuracy between SOA conditions, $F(1,57) = 5.95, MS_E = 0.040, p = .018$, letter set, $F(1,57) = 11.37, MS_E = 0.078, p = .001$, and response orders, $F(1,57) = 4.65, MS_E = 0.78, p = .035$. These significant main effects revealed that the 800 ms SOA condition (0.856) was more accurate than the 3200 ms condition (0.767); the same letter set (0.866) was more accurate than the different letter set (0.757); and mixed response order (0.897) was more accurate than blocked response order (0.726).

The ANOVA on proportion correct also revealed a marginal 3-way interaction, $F(1,57) = 2.87, MS_E = 0.040, p = .096$. Errors bars were calculated in accord with a method by Masson and Loftus (2003) and are shown in Figure 7, above. The Blocked-
Different condition was less accurate than all other condition in both SOA conditions and that the Mixed-Same condition was more accurate that all other conditions in the 3200 ms SOA condition.

*Response Time.* A mixed ANOVA on all conditions was also run with count-typing response time as the dependent variable. The ANOVA revealed that the blocked response order (2390 ms) was faster than the mixed response order (2850 ms), $F(1, 57) = 4.75, MSe = 1.32, p = .033$. No other results were significantly different.

*Signed Error.* As in Experiment 1, signed error was calculated by subtracting the correct answer from the participant’s response. A chi-square test was conducted on the frequencies of undercount and overcount errors. Table 2 represents the frequencies of undercounts and overcounts in each condition.

Table 2.
Frequency of undercount and overcount errors in Experiment 2.

<table>
<thead>
<tr>
<th>Response Order</th>
<th>Letter Set</th>
<th>SOA</th>
<th>Undercount</th>
<th>Overcount</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocked</td>
<td>Different</td>
<td>800</td>
<td>10</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Blocked</td>
<td>Different</td>
<td>3200</td>
<td>11</td>
<td>10</td>
<td>21</td>
</tr>
<tr>
<td>Blocked</td>
<td>Same</td>
<td>800</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Blocked</td>
<td>Same</td>
<td>3200</td>
<td>11</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>Mixed</td>
<td>Different</td>
<td>800</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Mixed</td>
<td>Different</td>
<td>3200</td>
<td>8</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Mixed</td>
<td>Same</td>
<td>800</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Mixed</td>
<td>Same</td>
<td>3200</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>54</td>
<td>31</td>
<td>85</td>
</tr>
</tbody>
</table>

*Note.* Only includes errors of five or less in magnitude.

All undercount and overcount errors were added across SOA conditions to determine the overall likelihood of each type of error. Overall, there was a 63.5 percent chance of an undercount and a 36.5 percent chance of an overcount. Each condition was tested individually to check if the frequency of undercounts and overcounts differed from
this likelihood proportion. No conditions differed significantly from the overall proportion.

*Error Monitoring.* Confidence ratings on a scale of 1 to 5 were divided into low confidence (ratings of 1 to 3) and high confidence (ratings of 4 and 5) for each condition. A proportion of low and high confidence was calculated for both error and correct trials. Figure 8 displays this proportion for error trials. Figure 9 presents this proportion for correct trials.

![Bar chart](image)

Figure 8. Proportion of error trials with low and high confidence as a function of SOA condition.
Figure 9. Proportion of correct trials with low and high confidence as a function of condition.

Figure 8 illustrates that when errors were made in the Blocked condition the participants showed a substantially greater chance of recognizing that the error was made, by rating the confidence in the count response as low. Figure 9 illustrates that when errors were not made, all conditions, except possibly the Blocked-Different-3200 ms condition were rated with equally high confidence.

*Letter-Typing Time.* After all load error trials were filtered out, only letter-typing time remained as a measure of letter-typing performance. A mixed ANOVA was run on all conditions of Experiment 2 and revealed significant differences between SOA, $F(1,57) = 5.30, MS_E = 2.31, p = .025$, between letter set, $F(1,57) = 25.27, MS_E = 15.57, p < .001$, and an interaction between SOA and letter set, $F(1,57) = 12.89, MS_E = 2.31, p <$
.001. No other significantly different results were found in letter-typing time. Figure 10 shows letter-typing time by SOA and letter set.

![Figure 10. Letter-Typing response time as a function of SOA and letter set condition.](image)

An analysis of error bars reveals that the different letter set, 800 ms SOA condition (9,184 ms) was slower than both the different letter set, 3200 ms SOA condition (7,560 ms) and the same letter, 800 ms SOA condition (4,598 ms). The different letter set, 3200 ms SOA condition was slower than the same letter.

**Discussion**

To summarize, the following are the important results from Experiment 2. First, accuracy was worse when the SOA was long in the conditions in which the participants had either a goal representation or verbal working memory load. Accuracy was also low when the letter set was different while the response order was blocked, whereas when neither condition was true, accuracy was very good. Although participants were more accurate in the mixed condition, they were also significantly slower in responding.
Second, no overcounting was discovered in any condition, even when the SOA was long. Another important result was that participants detect errors better when the errors are in context of a blocked response order. Finally, participants were slower when typing different letter sets than the same letter set.

The results show that there are disruptions to counting when under conditions of working memory load or long delays between items. The predominance of overcounting for long delays found in Experiment 1 was not replicated in this experiment, indicating the need to explain how the new procedure for Experiment 2 eliminated the effect from Experiment 1.

Delays between storing a representation and using the representation typically result in decay, in which the item loses strength over time. In the long delay of Experiment 1, decay of the prospective number will occur unless the participant uses a strategy to reduce the amount of decay. Mental rehearsal of the prospective number reduces the amount of decay and in Experiment 1 there was nothing to prevent the participant from rehearsing the number. In Experiment 2, a working memory load reduced the opportunity to rehearse the current total. As discussed in the Introduction, over-rehearsal of the current number may lead to more intention-outcome confusions by raising the strength of the prospective number to the point where it is confused for an outcome and the individual increments too many times. Reducing the amount of rehearsal should reduce the number of intention-outcome confusions.

Even in the conditions in which the verbal load is the same on every trial, participants may have chosen to rehearse the verbal load in memory to keep it active and accessible instead of over-rehearsing the total. The claim of the over-rehearsal
explanation is that with no working memory load, participants not only rehearse the prospective total more often, but rehearse it to the point that it is difficult to distinguish between the current and prospective totals. If participants in the same letter set, mixed response order, long SOA condition retrieve the letter set on some steps, then they may at once keep the same letter set fresh in memory, but also reduce the rehearsal of the prospective total just enough to reduce intention-outcome confusions. This theory has very little support in the data of Experiment 2, but data from the subsequent experiments without any working memory load demonstrate a tendency to overcount in long SOA conditions.

Undercounting with a working memory load is evidence against the generality of the theory by Buchner et al. (1998), which argues that when people forget they also overcount. Perhaps individuals do search for a number in memory when they have forgotten the prospective number, but do not overestimate after selecting the number, especially when a working memory load reduces the capacity necessary to execute such a strategy. I will expand on this explanation in the General Discussion.

Aside from examining the type of errors, Experiment 2 was designed to discover properties of goal representations. A preliminary aim was to confirm that the blocked conditions of the experiment actually encouraged participants to store an extra goal representation. The data indicate that the blocked condition was faster than the mixed condition, suggesting that participants in the blocked condition were ready to type the count response when the response screen appeared because they had been informed which request would be made first. The participants in the mixed condition needed to check the response cue to begin typing.
The counting data do not reveal any difference between storing an extra verbal load or an extra goal representation. Both conditions hurt performance, suggesting that both are stored in working memory, but not necessarily in different formats. Evidence from the confidence measure does reveal a difference in the effect of a goal representation load or a verbal load on the counting process. The data show that participants are more likely to be aware of an error when a goal representation load is stored than a verbal load. This effect cannot be completely explained by a general lower confidence when in the blocked condition because the number of low confidence responses is not much greater for the blocked, correct count trials than the mixed, correct count trials. These results suggest that storing an extra goal representation disrupts the flow of counting to the point where errors are more noticeable, whereas the same is not true of storing an extra verbal load. Carlson and Yaure (1990) demonstrated a similar phenomenon when they discovered that a competing intention disrupts processing more than short-term retention of information. I will examine how this disruption might occur in the General Discussion.

EXPERIMENT 3

The three hypotheses for why a long SOA disrupts counting are that increased decay causes more forgetting, increased rehearsal causes more intention-outcome confusions, and increased chance of distracters causes more premature triggering of intentions. Experiment 2 explored the forgetting and over-rehearsal hypotheses. Experiment 3 explored the distracter hypothesis.

The last piece of information necessary to complete a goal representation in a long SOA condition is the item. After completing the operation of the last intention, the
participant already knows the desired outcome (i.e., to calculate a new total) and the operator (i.e., to assign a new number). A simple retrieval from memory provides the prospective total relatively quickly and leaves only the item slot unfilled. A distracter mistaken as the target item may complete the intention erroneously. The goal of Experiment 3 was to manipulate the chance that a suitable distracter will appear. If distracters incorrectly trigger intentions, then more distracters will result in more overcounting.

Assuming a steady rate of infrequent distracters in the environment, Experiment 3 was designed to present distracters at a constant rate with a low probability (except within 100 ms before or 100 ms after the target events). In this experiment, target items (items that are to be counted) appeared within a box appearing in the middle of the screen. The participants were instructed to count only items that appeared inside the box, implying that all items appearing outside the box were not be to be counted (distracters).

Another hypothesis tested in Experiment 3 was that identical items increase ignoring of item labels and thus, susceptibility to distracters. In Experiment 3, half of the participants saw identical items and the other half saw distinctive items (each item was different from the two previous items).

Experiment 3 included an SOA factor. For each participant half of the trials had an 800 ms SOA and the other half had a 3200 ms SOA. The prediction was that the 3200 ms SOA condition should show more overcounting than the 800 ms SOA, irrespective of the distracters. Experiment 3 had no condition to prevent rehearsal of the prospective total and therefore should not have prevented overcounting in the long SOA condition.
Another manipulation was the probability that a distracter would appear. One third of the participants saw no distracters. The other two-thirds of participants saw distracters with two different probability rules. In the target-based distracter condition, the probability of a distracter was 0.1 for each target event that appeared. In the time-based distracter condition the probability of a distracter was 0.05 for each 200 ms of blank interval (excluding the first and last 100 ms of blank interval).

The prediction was an interaction in accuracy and overcounts between SOA and distracter group and an overall effect of item group (i.e., identical versus distinctive items). The long SOA condition was predicted to show lower accuracy and greater overcounting than the short SOA condition in the no distracter group (for the same reasons as Experiment 1), and even greater overcounting for the distracter groups, especially the time-based distracter group (which had the most distracters). Another prediction was an overall increase in performance when items were distinctive because the participant should be less likely to ignore the item identity information with distinctive items than identical items.

Method.

Participants. Ninety students in an introductory psychology course at The Pennsylvania State University participated in exchange for a small amount of course credit.

Design. The participants were divided into six groups of fifteen participants each. Three groups counted the same item in each trial and the other three groups counted distinctive items. Each of the item type groups counted in one of three distracter groups. One of these groups counted without distracters and the others counted with distracters.
One of the distracter groups counted with a probability of a distracter of 0.1 for each target that appeared, called the target-based group. The other distracter group counted with a probability of a distracter of 0.05 for every 200 ms of blank screen. Each participant had 10 trials of 800 and 3200 ms SOA for a total of 20 trials. Overall, the design was two between-subjects variables of distracter probability (zero, target-based, or time-based) and items (identical or distinctive) and one within-subjects variable of SOA (800 or 3200 ms).

Procedure. Except for the following changes, Experiments 1 and 3 will be identical. First, the number of items was a randomly chosen number between 25 and 50, as in Experiment 2. Participants were instructed to count only items that appeared inside a box that stayed on the screen throughout each trial even when the probability of a distracter was zero.

The box was a square, which appeared in the center of the screen with 168 pixel length sides. Every target item appeared in the center of the square. For the distinctive item groups, items were selected randomly from a pool of six Webding symbols (♥️ 🌚 🖔❗️➕🌟) on every step with the condition that no two items were the same within two steps of one another. For the identical item groups, all items were a randomly chosen Webding of the same set as the distinctive group chosen for each participant across all trials.

In the distracters groups, distracters appeared in a random location outside of the box within 10 to 84 pixels from the outer edge of the box. Thus, the distracters appeared in the area within half the length of one side of the square in any direction. Distracters appeared for 200 ms, the same length of time as a target.
In the distracter groups, the possibility of a distracter began 100 ms after the offset of the target asterisk. In the target-based distracter condition, the probability of a distracter was 10% for every target item that appeared and the distracter appeared within a randomly chosen 200 ms interval within the blank interval. In the time-based distracter condition, the probability of a distracter was 5% for every candidate 200 ms interval with the blank interval (two distracters could appear within one blank interval in this condition). No distracter appeared in the 100 ms interval preceding the next target. Overall, there was a 10% probability of one distracter for every target in both the 800 and 3200 ms SOA conditions in the target-based distracter condition. In the time-based distracter condition, there was a 10% probability of at least one distracter in the 800 ms SOA condition and a 75% probability of at least one distracter in the 3200 ms SOA condition.

Results

All responses of a magnitude greater than five from the correct response were excluded from all analyses. Errors of greater than five accounted for 1.2% of all data in this study.

Accuracy. Only responses that were the same as the correct total were considered accurate. All other responses were considered errors. Participants made correct responses on 88.0% of all trials.

A mixed analysis of variance (ANOVA) on the within-subjects levels of SOA (800 and 3200 ms) and the between-subjects levels of distracter probability conditions (zero, target-based, and time-based) and item type (identical and distinctive) demonstrated a reliable difference in accuracy between SOA conditions, $F(1,84) = 20.38,$
$MS_E = 0.011, p < .001$ with 800 ms SOA (0.913) more accurate than 3200 ms SOA (0.841). The ANOVA also revealed a marginal difference between distracter conditions. Figure 11 shows the distribution of proportion correct as a function of distracter conditions. The error bars indicate that the zero probability condition (0.917) was more accurate than the time-based probability condition (0.844).

Figure 11. Proportion correct as a function of distracter condition.

Signed Error. A mixed ANOVA was run on signed error and revealed a significant difference between SOA conditions, $F(1,84) = 7.88, MS_E = 0.047, p = .006$. The 800 ms SOA condition showed more undercounting (-0.065) and the 3200 ms SOA condition showed more overcounting (0.026).

As in Experiments 1 and 2, a chi-square test was conducted on the frequencies of undercount and overcount errors. Table 3 represents the frequencies of undercounts and overcounts in each condition.
Table 3.

Frequency of undercount and overcount errors in Experiment 3.

<table>
<thead>
<tr>
<th>Distracter Probability</th>
<th>Item Type</th>
<th>SOA</th>
<th>Undercount</th>
<th>Overcount</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero</td>
<td>Distinctive</td>
<td>800</td>
<td>6</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Zero</td>
<td>Distinctive</td>
<td>3200</td>
<td>10</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Zero</td>
<td>Identical</td>
<td>800</td>
<td>12</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Zero</td>
<td>Identical</td>
<td>3200</td>
<td>6</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Distinctive</td>
<td>800</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Distinctive</td>
<td>3200</td>
<td>5</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Identical</td>
<td>800</td>
<td>14</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Identical</td>
<td>3200</td>
<td>15</td>
<td>19</td>
<td>34</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Distinctive</td>
<td>800</td>
<td>7</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Distinctive</td>
<td>3200</td>
<td>16</td>
<td>14</td>
<td>30</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Identical</td>
<td>800</td>
<td>17</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Identical</td>
<td>3200</td>
<td>15</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>128</td>
<td>86</td>
<td>214</td>
</tr>
</tbody>
</table>

Note. Only includes errors of five or less in magnitude.

All undercount and overcount errors were added across SOA conditions to determine the overall likelihood of each type of error. Overall, there was a 59.8 percent chance of an undercount and a 40.2 percent chance of an overcount. Each SOA condition was tested individually to check if the frequency of undercounts and overcounts differed from this likelihood proportion. Only the Time-Based, Distinctive Item, 3200 ms condition demonstrating a greater number of overcount errors and a lower number of undercount errors than predicted $\chi^2 (1) = 6.54, p < .05$.

Error Monitoring. Confidence ratings on a scale of 1 to 5 were divided into low confidence (ratings of 1 to 3) and high confidence (ratings of 4 and 5) for each condition. A proportion of low and high confidence was calculated for both error and correct trials. There was no discernible pattern from this data.
Discussion

Experiment 3 did not reveal many statistically significant effects. The most significant findings were that the long SOA condition was less accurate and showed more overcount errors than the short SOA condition. Also, the time-based distracter condition was less accurate than the no distracter condition. However, there was very little support for more overcounts among the distracter conditions.

There was very little difference between the distracter conditions. Only the time-based distracter condition was significantly lower than the no distracter condition, suggesting that distracters are largely ineffective in hurting counting performance. In addition, a complete lack of effect of item type on performance indicates that counting is unaffected by identical versus distinctive items. Overall, the data indicate that participants generally succeeded in counting only items that appeared inside the box. Perhaps the distracter manipulation was not strong enough to counter the participant’s ability to focus inside the box.

Making the distracters manipulation stronger might have been achieved by adding more distracters, but a higher probability of distracters would not have modeled the hypothesis that distracters occur infrequently in the natural environment. If distracters are simply ignored in this experiment, then a new manipulation should be used in which participants cannot ignore distracters, but should not count them. This was one of the reasons for conducting Experiment 4.

Another result was that only the 3200 ms SOA condition and specifically the target-based, distinctive item, 3200 ms SOA condition showed more overcounting than undercounting. The long SOA overcounting effect was a replication of the same effect
from Experiment 1 and was expected because of the opportunity to rehearse in Experiment 3. However, the minimal effect of distracters on overcounting needs to be explained.

When the probability of a distracter was a function of time, there was a corresponding lower accuracy. This finding shows that there is a disruption of counting with distracters under the right conditions, but not for a specific type of error. There was no evidence that distracters only cause overcounts. When a distracter appears on the screen it may attract attention, but not trigger the intention to count. If the distracter attracts attention away from the target, then the participant may miss a target that should have been counted and undercount the targets. If triggering of intentions is prevented by simply ignoring the distracters, Experiment 4 should reveal more about the effect of distracters on counting when the individual is no longer able to ignore distracters.

EXPERIMENT 4

Experiment 2 focused on the forgetting and over-rehearsal theories of disruptions at long SOAs and Experiment 3 focused on the distracter theory. The aim of Experiment 4 was to focus on potential interactions among forgetting, over-rehearsal, and distracters.

There were three groups of participants. One group counted target asterisks and pressed a button for every target that they counted. Another group counted target asterisks and pressed a button for every distracter. The third group counted targets while holding down the button and served as a control group. Under this design, the nature of the goal representation load was somewhat different from Experiment 2. In Experiment 4, the loaded goal representation was to press the button and was enacted throughout the
In Experiment 2, the loaded goal representation was for a future action (i.e., to type the letter load) and was only enacted once at the end of the trial.

According to the forgetting hypothesis, groups with a goal representation load (to press the button) should show lower performance than the control group. The load should also interfere with the participant’s ability to rehearse the prospective number and therefore should reduce the number of intention-outcome confusions, which result in overcount errors. The prediction was that participants who press for each distracter would make more overcount errors because distracters cannot be ignored and would be more likely to trigger the intention to count.

To equate the two groups on the number of key presses made, a distracter appeared before every target. The distracter-press participants were required to make no fewer key presses than the target-press group. Another potential difference between the two press groups was the nature of the button-press timing. If targets appeared rhythmically, but distracters appeared with varied timing (as in the previous experiments), then the target-pressing group would benefit from the rhythmic pressing of the button. To equate the two groups, both targets and distracters appeared with varied timing.

In addition, Experiment 4 included the reading span measure by Daneman and Carpenter (1980). The reading span measure was designed to calculate the working memory capacity of individuals. The reading span measure is a working memory capacity measure because it calculates short-term memory storage while individuals engage in controlled processing (i.e., reading). Working memory is an important
construct in the theorizing above and the reading span measure helps to distinguish the role of the size of an individual’s working memory capacity.

Method.

Participants. Fifty-five students from an introductory psychology course at The Pennsylvania State University participated in this study in exchange for a small amount of course credit. One participant was removed from further analyses for going over the maximum one hour for the experiment. Three participants were removed from further analyses for being non-native English speakers (all participants were required to be native English speakers for the working memory span task). Six participants were removed for being left-handed (all participants were required to be right-handed for the button-pressing task).

Design. The remaining participants were divided into three groups of fifteen participants each. In the first group, participants were instructed to press the left mouse button every time a target appeared. The second group was instructed to press the same button whenever a distracter appeared. The third group was instructed to hold down the button throughout each trial. SOA varied within-subjects with 10 trials each of 800 and 3200 ms SOA. All participants received the reading span measure (Daneman & Carpenter, 1980).

Materials. Additional materials used in this experiment were created to administer the reading span measure. Prewritten sentences were printed, trimmed, and glued separately to colored, poster board cards. The cards consisted of five groups of two, three, four, and five sized sentence sets. The color of the poster board was color coordinated by sentence size to aid experimenters.
Procedure. The procedure of Experiment 4 was identical to that of the identical items-distracters group in Experiment 3 (i.e., using asterisks instead of Webding symbols), with the following exceptions. Instead of the different distracter probability values, a distracter appeared before every target at a randomly chosen 200 ms blank interval. The timing of item presentation was also different from Experiment 3. Instead of rhythmic SOAs (e.g., 3200 ms on every step), SOA was selected randomly from a pool of numbers on each step. In the 800 ms condition, the pool of SOAs was – 650, 750, 850, and 950 ms. To accommodate the possible SOAs and the condition that distracters appear in a 200 ms interval, the 125 ms before and after each target was excluded from the possibility of a distracter instead of the 100 ms before and after each target. In the 3200 ms condition, the pool of SOAs was – 2600, 3000, 3400, and 3800. In both cases the average SOA was the same as the other experiments. The only other change to the counting task was that the feedback screen also included accuracy and error feedback on the number of times the button was pressed (only for the groups that press the button for distracters or targets).

All participants were administered the reading span measure, created by Daneman and Carpenter (1980) and designed to measure an individual’s working memory capacity. Participants were instructed to read a series of sentences, presented by the experimenter, and to recall the last word from each sentence in a recall phase after reciting all sentences. The experimenter presented each sentence with a thumb over the last word in the sentence to prevent participants from looking ahead. The experimenter put down a new card after the participants spoke the last word of the sentence and flipped over all cards after the last word of the last sentence while asking for recall. Participants were
given a score equal to the last set in which they named all of the last words in the sentence, in all but one of the sentence sets. The participants received an extra half point for recalling all the words in all but two of the sentence sets.

**Results**

All responses of a magnitude greater than five from the correct response were excluded from all analyses. As in Experiment 1, these responses were considered as either typing errors or misremembering the ten’s digit. Errors of greater than five accounted for 2.8% of all data in this study.

**Accuracy.** Only responses that were the same as the correct total were considered accurate. All other responses were considered errors. Participants made correct responses on 80.8% of all trials. Figure 12 shows the distribution of proportion correct as a function of SOA and secondary task conditions.

Figure 12. Proportion correct as a function of SOA and Press Instruction conditions.
A mixed analysis of variance (ANOVA) on the within-subjects levels of SOA (800 and 3200 ms) and the between-subjects levels of secondary task (no press, distracter-press, and target-press) demonstrated a reliable difference in accuracy between secondary task conditions, $F(2,42) = 14.49$, $MS_E = 0.040$, $p < .001$ and a significant interaction between SOA and secondary task, $F(2,42) = 3.86$, $MS_E = 0.042$, $p = .028$.

Accuracy for distracter-press in both SOA conditions (0.645) was lower than both no press (0.912) and target-press (0.846) and that distracter press-800 ms SOA (0.569) was lower than distracter press-3200 ms SOA (0.721).

Signed Error. As in the previous experiments, a chi-square test was conducted on the frequencies of undercount and overcount errors. Table 4 represents the frequencies of undercounts and overcounts in each condition.

Table 4.
Frequency of undercount and overcount errors in Experiment 4.

<table>
<thead>
<tr>
<th>Press Instructions</th>
<th>SOA</th>
<th>Undercount</th>
<th>Overcount</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>800</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>None</td>
<td>3200</td>
<td>11</td>
<td>12</td>
<td>23</td>
</tr>
<tr>
<td>Target</td>
<td>800</td>
<td>16</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Target</td>
<td>3200</td>
<td>16</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Distracter</td>
<td>800</td>
<td>39</td>
<td>19</td>
<td>58</td>
</tr>
<tr>
<td>Distracter</td>
<td>3200</td>
<td>18</td>
<td>23</td>
<td>41</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>102</td>
<td>66</td>
<td>168</td>
</tr>
</tbody>
</table>

Note. Only includes errors of five or less in magnitude.

All undercount and overcount errors were added across SOA conditions to determine the overall likelihood of each type of error. Overall, there was a 60.7 percent chance of an undercount and a 39.3 percent chance of an overcount. Each condition was tested individually to check if the frequency of undercounts and overcounts differed from this likelihood proportion. Only the distracter-press, 3200 ms SOA condition showed a
significantly greater frequency of overcounts than the overall probability, \( \chi^2 (1) = 4.86, p < .05 \).

Although the frequency of overcounts in the distracter-press, long SOA condition does not exceed the frequency of undercounts by a large proportion (five more overcounts than undercounts out of 41 total errors), this condition showed a shift in the direction of overcounts. This finding indicates that additional errors may be made during counting that either compensate for other errors (e.g., an error leading to an overcount may compensate for a previous error that lead to an undercount) or lead to incorrect responses at the end of the sequence. The overall proportion of undercounts and overcounts was chosen as a better referent for each individual condition for direction of errors within the counting sequence.

*Error Monitoring.* Confidence ratings on a scale of 1 to 5 were divided into low confidence (ratings of 1 to 3) and high confidence (ratings of 4 and 5) for each condition. A proportion of low and high confidence was calculated for both error and correct trials. Figure 13 displays this proportion for error trials. Figure 14 presents this proportion for correct trials.
Figure 13. Proportion of error trials with low and high confidence as a function of Press Instruction-SOA condition.

Figure 14. Proportion of correct trials with low and high confidence as a function of Press Instruction SOA condition.
Figure 13 illustrates that when errors were made in the target-press and distracter-press conditions the participants showed a substantially greater chance of recognizing that the error was made, by rating the confidence in the count response as low. Figure 14 illustrates that when errors were not made, all conditions, except the Distracter Press conditions were rated with equally high confidence.

**Press Accuracy.** Overall, participants in the Distracter Press and Target Press groups pressed the correct number of times on 28.3% of trials. A mixed ANOVA on secondary task (distracter and target press) and SOA (800 and 3200 ms) demonstrated a difference between press accuracy for SOA, $F(1,28) = 18.64$, $MS_E = 0.015$, $p < .001$. The 800 ms SOA condition (0.351) was more accurate than the 3200 ms SOA condition (0.215). No other results were significantly different.

**Absolute Press Error.** A mixed-factor ANOVA was run on the median absolute error of presses to analyze secondary task performance beyond the overall accuracy. The data revealed a large range of press errors so, median press error reflected the best measure of central tendency. The ANOVA revealed a significant main effect of SOA, $F(1,28) = 11.64$, $MS_E = 2.35$, $p = .002$ and a significant interaction between SOA and secondary task, $F(1,28) = 16.54$, $MS_E = 2.35$, $p = .001$. Figure 15 illustrates the interaction between SOA and the secondary task. An analysis of errors bars reveals that the Target-Press, 3200 ms SOA (3.90) was larger than both Target-Press, 800 ms (1.50) and Distracter-Press, 3200 ms (2.23).
Figure 15. Median absolute error of presses from correct for each secondary task and SOA condition.

**Signed Press Error.** A mixed-factor ANOVA was run on the median signed error of presses. The ANOVA revealed a significant main effect of SOA, $F(1,28) = 11.64$, $MS_E = 3.46$, $p = .002$ in which the 3200 ms SOA condition (2.25) showed more overpressing than the 800 ms SOA condition (0.55). No other statistically significant differences were found.

**Working Memory Task.** The mean working memory span score across participants was 2.66 with a standard deviation of 0.601. A correlation was calculated for the average accuracy of trials for all subjects and working memory span score. The correlation was 0.135. A similar correlation between span score and absolute errors of counting was -0.128.
Discussion

Several results of Experiment 4 stand out. First, when the secondary task was to press for each distracter, accuracy was lower than each of the other two groups, whereas pressing for each target did not result in lower accuracy than not pressing. Second, the increase in errors for the distracter press task at 3200 ms was actually an increase in overcounting errors. Another important result was that error monitoring was better for the conditions that had a secondary task. Secondary task performance was worse for the 3200 ms SOA condition, which showed more overcounting than the 800 ms SOA condition and absolute error of pressing was greatest for the target-press, 3200 ms SOA condition. Finally, working memory span scores correlated positively with counting accuracy and negatively with absolute counting error, but neither span score was particularly large.

A reason for conducting Experiment 4 was to examine counting when participants are unable to ignore distracters. The data indicate that counting performance decreases more when participants are instructed to press a button for each distracter than when pressing for each target. Not only is pressing for distracters disruptive, but it also increases the number of overcounts in the 3200 ms SOA condition. There is an increase in overcounts even though participants store an extra goal representation (i.e., to press the button). Experiment 2 showed that storing an extra goal representation increased the number of undercounts. Together, these results suggest that distracters can trigger intentions prematurely, so long as participants cannot actively ignore the distracting area.

The distracter-press findings also have implications for the nature of goal representations. Distracters and targets are distinct from one another in both their spatial
location (i.e., outside versus inside of the box) and their temporal appearance (i.e.,
distracters do not appear at the same time as targets). When pressing for distracters, two
distinct goal representations need to be constructed and executed; one for counting targets
and one for pressing when a distracter appears. When pressing for targets, two goal
representations may need to be constructed, but those goal representations can be linked
along the common element – the target item. In other words, executing one intention
may trigger the next intention. Linking the two goal representations together triggers
more complex processing when the item appears, but once both intentions are complete,
the participant does not need to store the extra goal representation in an active state,
because completing the first intention triggers the next intention.

Further evidence supports the linked goal representation theory. Pressing was
more difficult in the 3200 ms SOA condition (in overall accuracy) and especially the
Target-Press, 3200 ms SOA condition (in absolute error). This result suggests that
participants had a more difficult time remembering pressing when the SOA was longer
because of the need to remember over a longer time while decay reduces the strength of
the memory for pressing. The finding that participants generally over-pressed in the
Target-Press, 3200 ms SOA condition supported the theory that people forget whether
they pressed. If participants press too many times then they have likely forgotten that
they pressed and press again to compensate. If both pressing and counting are in linked
intentions, then the memory of pressing will experience interference with the primary
task of counting. I will return to this topic in the General Discussion.
GENERAL DISCUSSION

The purpose of these studies was to understand goal representations, especially how they are stored and how they are used. The studies were conducted with a focus on how people process a task with multiple steps (counting) by using goal representations. Experiment 1 revealed that individuals tend to count too many times when processing events at a long SOA. Experiment 2 revealed that this increase in overcount errors was not necessarily caused by forgetting, but may have been caused by rehearsing too many times. Experiment 2 also showed that when people forget because of a goal representation load, they are also more likely to notice the error, more so than when they forget because of a verbal load.

Experiment 3 focused on distracters as a potential cause of overcounting. According to the distracters hypothesis, distracters may be mistaken as targets and trigger an intention prematurely. Although Experiment 3 did not reveal convincing evidence that distracters trigger intentions, Experiment 4 showed that when participants are unable to ignore distracters, distracters may trigger intentions to count. Experiment 4 also revealed that an extra goal representation with a common property may also be linked together or integrated.

The following sections discuss different implications of these and other findings. The discussion begins with explanations of overcounting and moves onto discussions about the storage of goal representations, triggering conditions of intentions, and how goal representations may be linked. I will then discuss other important considerations, limitations of the studies, and future directions of this research.
Explanations of Overcounting

Overestimation Strategy Account. One potential explanation of overcounting was a theory proposed by Buchner et al. (1998). Buchner et al. believed that counting errors in their own experiments were the result of forgetting the current number and using a strategy to retrieve the number. According to the theory, when people notice that they have forgotten the number to be incremented, they search their memory and select the most active number. The selected number is a starting point as the individual selects a higher number for the current total. The number originally selected from memory was most likely a number that had been activated in the past and the participant has no efficient method for discriminating when it had been activated. Selecting a higher number may be the best approach to choosing the correct number.

Buchner et al.’s theory assumes that overcounting is a result of forgetting and a long SOA increases the amount of decay, a major source of forgetting. Experiment 1 demonstrated overcounting with a long SOA. Experiment 2 was designed to increase the amount of forgetting through working memory loads. Experiment 2 demonstrated disruption in counting from the memory load conditions (i.e., lower accuracy), but did not show the expected overcounting effect. The lack of overcounting was disconfirming evidence for Buchner et al.’s strategy explanation.

Perhaps the overestimation strategy explanation was partially correct. Instead of forgetting, searching for a number in memory, and then over-estimating, participants searched for the number in memory and used that number as the current number. This explanation would predict undercounting (i.e., a number drawn from memory is likely to
be one that has been activated prior to the current step), yet not rely on the claim that all individuals use the same complex strategy of overestimation.

*Over-Rehearsal Account.* If forgetting does not explain overcounting with a long SOA, then perhaps intention-outcome confusions can explain the phenomenon. Intention-outcome confusions occur when the individual mistakes a prospective number contained in the goal representation for the outcome of counting. When participants commit an intention-outcome confusion they will overcount because they will use a number that has been incremented without a corresponding target item.

One potential cause of intention-outcome confusions is that the prospective number becomes so active in memory that it is confused as the number in working memory representing the current total. Increased activation of the prospective number is most likely to occur in a long SOA condition. Items in short-term memory decay or weaken over time (e.g., Peterson & Peterson, 1959); the more time that passes the weaker items become. Forgetting is most likely to occur in a long SOA condition without the intervention of a strategy.

The most effective strategy for keeping a short-term memory object active is to mentally rehearse the object or keep it strong by repeating it to oneself (Atkinson & Shiffrin, 1968). To counteract the decay problem of long SOAs, participants are likely to rehearse the current number more in the long SOA condition, leading to intention-outcome confusions and thus, overcounts.

The data across Experiments 1 through 4 support the rehearsal explanation of overcounting. In Experiment 1, nothing prevented rehearsal and therefore the Long SOA condition demonstrated overcounting. In Experiment 2, working memory loads prevent
excess rehearsal. Even in the mixed response order, same letter set, long SOA condition, participants could choose to devote some working memory capacity to the letter load, preventing over-rehearsal. In Experiment 3 the long SOA condition shows more overcounting than the short SOA condition. In Experiment 4, the long SOA conditions always have an extra intention (i.e., pressing) except the no press condition (which showed more overcounts than undercounts, it just was not significantly different from the other conditions).

The finding that more rehearsal leads to worse performance is surprising. When people mentally rehearse, they are attempting to strengthen memory, but in the case of counting in which the running total must be constantly updated (see Altmann & Gray, 2002), rehearsal leads to unwanted results. Until now, researchers have only found that rehearsal helps performance (e.g., Atkinson & Shiffrin, 1968) or does not affect performance (e.g., Craik and Watkins, 1973).

*Distracters Account.* The distracters hypothesis is also a plausible explanation that can explain overcounting at long SOAs. According to the distracters hypothesis, individuals count distracters that appear in the environment (e.g., a light flicker) because distracters are mistaken for targets and trigger the intention to count.

Experiment 3 explored the distracters explanation. Distracters did not have much of an effect on counting. The condition with the most distracters (i.e., the time-based distracter condition) was the only distracter condition that decreased counting accuracy and only the time-based, distinctive item, 3200 ms SOA condition revealed a predominance of overcount errors. However, this predominance of overcounting may be attributed to the predominance of overcounting in the long SOA condition.
The evidence of overcounting in the long SOA condition was expected. The memory load from Experiment 2 was not present in Experiment 3 and so overall; the results of Experiment 3 should match the results of Experiment 1. The lack of overcounting from the distracters conditions disconfirms the distracters theory of overcounting. Clearly, distracters do not cause overcounting by themselves in the context of Experiment 3. If participants are instructed to not count distracters as they were in Experiment 3 and not Experiment 1, then they are effective at avoiding triggering incorrect intentions. Experiment 3 provides no evidence as to whether distracters pull attention away from targets (causing undercounts) or incorrectly trigger intentions to count (causing overcounts). Either way, the distracter manipulation was not a very strong effect in Experiment 3.

One of the reasons for Experiment 4 was to ensure that participants could not ignore the distracters, by instructing a selection of participants to press a button every time a distracter appeared on the screen. A potential cause of the weakness of the distracter manipulation in Experiment 3 was that there was no reason to attend to the distracters, so participants were free to ignore the area outside the box.

Pressing the button for every distracter lowered count accuracy and caused overcounting in the 3200 ms SOA condition. In the distracter press 800 ms SOA condition counting errors were largely undercounts. The difference in type of error between distracter press, 800 and 3200 ms SOA, illustrates the above conclusion that distracters can either incorrectly trigger intentions or move attention away from a target item. In the 800 ms condition, a distracter occurs right before the next target and if focus is moved to the distracter in order to press the button, then the target may not be counted.
In the 3200 ms condition, the distracter is less likely to occur just before the target, but may cause inadvertent triggering of the next intention to count. This account explains the split between tendency to overcount and undercount in the SOA conditions of the distracter press group.

Participants may have ignored distracters in Experiment 3, but there is no accurate method for discriminating whether or not participants also ignored natural distracters in Experiment 1. Experiment 4 revealed that distracters that require attention are likely to cause overcounts, perhaps from incorrectly triggering intentions to count. Distracters may or may not trigger intentions, but certainly they sometimes cause participants to miss targets or fail to increment, as in the distracter-press group of Experiment 4 in the 800 ms SOA condition. Follow-up studies will need to be run to check this possibility.

The Nature of Goal Representations

Storage of Goal Representations. In Baddeley and Hitch’s (1974; Baddeley, 1986) model of working memory, the three components of working memory are the central executive for controlling information, the visuo-spatial sketchpad for storing visual and spatial information, and the phonological loop for storing acoustic information. Baddeley and Hitch (1974) settled on this conclusion after finding that individuals can store near capacity visual information and phonological information at the same time (see Baddeley, 1986 for a review). Baddeley and Hitch hypothesized that there must be separate working memory stores for each type of information to achieve this result.

Another possible source of the ability to store near capacity visual and phonological information is a lack of interference, a major cause of forgetting (Logan
2004; Waugh & Norman, 1965), if not the primary cause of forgetting (Nairne, 2002). Perhaps visual and phonological information are stored with distinct formats that are not confused with one another. For example, several words with similar sounds are much more difficult to store than words with distinctive sounds (e.g., Baddeley, 1966) because the similar words can be confused with one another and use more attentional capacity to remember them. Format-based interference is the perspective I will assume, rather than the componential approach to working memory.

According to Baddeley and Hitch (1974), the central executive is also distinct from the other components of working memory. They suggest that the central executive’s function is to control the flow of information, to store some task-relevant information, and to process complex tasks. Executive-control demanding tasks, such as random number generation (e.g., Gilhooly, Logie, Wetherick, & Wynn, 1993) interfere with other executive tasks, but not visual or phonological memory capacity (see Baddeley, 1986). Perhaps the component known as the central executive is actually a distinctive format of information in working memory rather than a distinct component.

If the purpose of the central executive is cognitive control, then the best candidate format of information in the central executive is a goal representation format. The purpose of a goal representation is to store task-related instructions for working memory. Perhaps studies showing that central executive tasks do not interfere with storing visual and phonological information indicate that goal representations are stored in distinctive formats from visual and phonological information, which do not interfere with one another. Logan (2004) addresses this issue with his task span procedure, asserting that memory for goals is different from memory for other types of information.
Evidence for this perspective was not found in counting performance measures of the above experiments. Accuracy results in Experiment 2 showed worse counting performance when there was either a verbal load or a goal representation load. This evidence is in favor of both types of load hurting performance, but there is little difference in how they affect counting itself.

The data addressing a difference between the two types of load comes from the confidence measures. In Experiment 2, participants showed higher confidence when they made an error (and therefore should have had low confidence in their response) in the condition without an extra goal representation load than the condition with an extra goal representation. According to Carlson and Cassenti (in press), high confidence after an error is indicative of fluid processing. Error detection is triggered by a disruption to fluent counting (in the case of Carlson & Cassenti, in press, the disruption was non-rhythmic timing of events).

If individuals notice errors more when there is a disruption, then storing an extra goal representation must be more disruptive than any other condition in Experiment 2. The same may be said of the press conditions of Experiment 4, in which participants have more low confidence responses when committing errors than the no press group. I propose that this disruption is caused by forgetting the goal representation for counting and needing to reconstruct it, as opposed to merely forgetting the current total and selecting another from memory. A disruption requiring reforming a complex representation is disruptive in time and a noticeable interruption of fluid, controlled processing (see Carlson & Yaure, 1990).
Instead of a working memory with several components, I assume that working
memory is a set of mental processes that combine computational abilities and storage of
several types of representation. These representation types are stored in the same short-
term stores and consist of at least phonological, visual, and goal representations. More
load of a certain type of representation will only cause interference of other stored
representation with the same type of format (see Nairne, 2002).

Interference is a concept adopted by Baddeley and Logie (1999), who assume that
similarity-based interference within the visuo-spatial sketchpad and phonological loop is
one of the factors responsible for capacity limits of visual and auditory representations.
For example, forgetting within the phonological loop is increased when words with
similar sounds (e.g., hat, bat, and cat) are stored rather than words with distinct sounds
(Baddeley & Hitch, 1994; Conrad and Hull, 1964). The perspective adopted here
suggests that if only similar representations interfere with one another, then all types of
representations may be stored in the same short-term storage and still demonstrate
decreased performance when too many representations of the same format (which are
more similar than different format representation) are used at once. Baddeley’s (1986)
assumption of separate working memory storage components is therefore unnecessary
(see Cowan, 1999). Indeed visual and verbal representations may interfere with one
another under the right conditions (Kessels & Postma, 2002), an improbable situation if
the representations are stored in separate workspaces.

Executive control (a working memory function ascribed to Baddeley’s, 1986,
third component of working memory, the central executive) may have its own distinct
form of representation, the goal representation (see Logan, 2004; Logan and Bundeson,
Experiment 2 showed that a goal representation load can decrease task performance by interfering with the goal to count and disrupting the flow of processing. Only the goal representation load condition demonstrated an increase in error detection, indicating a differential effect of goal representation load and verbal load on performance. These distinct patterns of performance implicate a difference between verbal and goal representations. These results support the representational perspective of executive control and a distinct format of goal and verbal representations questions the assumption of a separate component of the central executive.

Including all working memory storage into one system of short-term storage contrasts with the Baddeley (1986) model of working memory, but is consistent with many other models of working memory (see Miyake & Shah, 1999, for a review of several models of working memory). For example, the Cowan (1999) model of working memory assumes that both visual and auditory information is stored in the same short-term memory buffer and that interference between representations of the same type accounts for capacity limits.

The claim that goal representations are the format of representation for the executive control functions of working memory is the assumption that differs from other models of working memory. Prominent theories of working memory, including Cowan’s (1999, 1988) and Baddeley’s (1986: Baddeley & Logie, 1999) assume a separate component for the central executive. However, Baddeley and Logie (1999) point out that the central executive is the component of working memory that is the least understood. Incorporating the goal representation into a theory of executive control poses research questions that may help de-mystify the central executive.
Linking goal representations. Goal representations are a special type of working memory representation because there is some measure of control in how people represent goals. Research into phonological representations indicates that phonological representations are stored in a similar way as the utterances are produced. For instance, research indicates that the time to mentally rehearse a to-be-remembered word is proportional to the time to say the word aloud (e.g., hippopotamus takes longer to say than dog, Baddeley, Thomson, & Buchanan, 1975). Similarly, mental, visual images possess the same properties as pictures in the environment (e.g., mental rotation time, Shepard & Metzler, 1971).

Goal representations, however, have some flexibility in how they are represented. Consider the situation in Experiment 4, in which some participants must press the button when they see a distracter and some must press the button when they see (and count) a target. In the target press condition, two goal representations may be represented: “I intend to assign the next number to the next item” and “I intend to press this button when I see the next item.” An alternative approach is to form two linked intentions: “I intend to assign the next number to the next item, then I intend to press this button” or two integrated intentions: “I intend to assign the next number to the next item and to press this button.” The advantage of the distinct goal representations approach is that set-up is less complex without the extra link (i.e., “then”). The advantage of the linked or integrated goal representation approach is that one of the two goal representations triggers the other, reducing the need to store both representations in working memory (i.e., store the second intention in long-term memory).
The results of Experiment 4 shed light on how participants construct goal representations. The target press group was no different from the no press group in accuracy for either SOA condition, but was better than the distracter press group in both SOA conditions. Clearly, the target press group did not suffer from the same decrement in performance as the distracter press group.

Experiment 2 showed that storing an extra goal representation hurt counting performance, but Experiment 4 did not show a similar decline even with an extra goal representation to press when a target appeared. These findings indicate that participants may avoid taxing working memory when using two intentions linked together with a common element (e.g., the to-be-counted item) rather than two intentions with no common elements (e.g., counting targets and pressing for distracters).

Linking the two intentions may have benefited counting, but linking also may have hurt pressing performance. Results of Experiment 4 showed a higher absolute error in the target-press, long SOA condition. If the intention to press and the intention to count were linked, then participants may have had a difficult time remembering if they had done both tasks. The memory is worse in the long SOA condition because memory decay reduces strength of the episodic memory for pressing.

The memory of pressing would be worse if the intention to count triggers the intention to press because participants would not have constructed the pressing intention independent of the counting intention. Instead, participants would have relied on the link between the two intentions. If the participants forgot that they had pressed already, then they would have pressed again to compensate for a missed press. The results confirm this reasoning as participants tended to over-press rather than under-press. The problem of
over-pressing is confounded by the fact that correcting for a detected over-press is less obvious than correcting for a detected overcount. In the former case, participants would need to wait until the next trial and not press to compensate, whereas in the latter case, participants can simply decrease their current total by one.

Triggering of intentions. Another implication on the nature of intentions is how individuals trigger intentions. According to the literature on stop-signals (e.g., Logan, 1983), individuals often have a difficult time stopping an intention that has already been constructed. One of the primary goals of Experiment 3 was to discover if intentions could be triggered by events that were similar to target events.

In Experiment 3, some participants saw distracters that they were instructed not to count. The distracters always appeared outside a box that was in the middle of the screen, but were always of the same item type as the target events, which appeared inside the box. If intentions triggered from distracters, then participants should have made more overcount errors when distracters appeared.

There was no indication that distracters trigger intentions in this task. There was a disruption from distracters in one of the conditions, but errors showed no trend towards overcounting. This does not preclude that distracters sometimes trigger intentions, but suggests that participants are good at ignoring distracters when instructed that distracters exist. Perhaps in Experiment 1 natural distracters did cause overcounting because participants were not instructed about their existence.

If avoiding overcounts comes from instructions that encourage ignoring distracters, then overcounting may be found in conditions in which participants cannot ignore distracters. This was the case in the distracter press group of Experiment 4.
Participants in the distracter press group were asked to press a button every time a distracter appeared and were later given feedback about their pressing performance. Therefore, participants could not ignore the distracters while following the instructions.

The results of Experiment 4 showed that distracters cause more overcounts when the participant directs attention to them at least in the long SOA condition. The tendency towards overcounts appeared despite the goal representation load, which tended to cause undercounts in Experiment 2. This result implies that distracters can cause incorrect triggering of intentions, so long as the focus of attention moves to the distracters.

The distracter press group from Experiment 4 is the condition that most closely matches the procedure of the stop signal studies. In both cases, participants are inclined to focus on the presented stimulus, but must inhibit responding to it. In Experiment 4, participants must focus on distracters in order to press the button the correct number of times. In most trials of the stop signal studies, participants need to respond to the stimuli as quickly as possible; stop signal trials were rare.

If individuals need to focus on the distracter to incorrectly trigger an intention, there is no telling whether or not the overcounting at a long SOA in Experiment 1 was the result of natural environmental distracters. The instructions of Experiment 1 did not indicate anything about distracters. If a natural distracter had appeared, the participants may or may not have incorrectly counted it.

*Other Considerations*

*Individual differences of working memory capacity.* Individuals vary in working memory capacity or how many items they can represent at one time (e.g., Kane et al., 2001). Therefore, individuals with a higher working memory capacity should perform
better when counting with a working memory load because their working memory capacity will allow them to store both the load and the count at the same time. All participants in Experiment 4 completed a working memory capacity measure (Daneman & Carpenter, 1980) and their measured capacity was compared with their counting accuracy and absolute error.

The results showed a trend for higher accuracy and lower absolute error when working memory capacity was high, but this trend was not especially large. Two explanations may account for this finding. The first is that working memory capacity only negligibly affects counting performance and there were not enough participants or trials to provide the power to discover a difference in performance between different sizes. Perhaps more trials with more participants would produce stronger finding.

The second potential explanation is that the working memory test, designed to test capacity for verbal material, was not the correct tool for detecting differences in capacity for goal representations. In Experiment 2, participants detected errors better when storing an extra goal representation. This finding suggests that storing a goal representation load is a different type of phenomenon than storing a verbal load. Assuming different types of representations, perhaps a new type of working memory test, which focuses on storage of goal representations will reveal a significant trend for higher working memory capacity and better counting performance (see Logan, 2004, for progress towards this goal).

The data from Experiment 4, reveals very little about which of these two explanation is the correct one or if both explanations explain portions of this finding. Follow-up experiments will need to be conducted to figure out which explanation is more accurate.
The representation of numbers. People can represent numbers in one of two ways, verbally or analogically. Verbal representation of numbers is in a pronounceable format, such as “one” or “two” (see Logie, & Baddeley, 1987; Buchner et al., 1998; Carlson, & Cassenti, in press). Analogical representations are non-pronounceable representations of quantity, which acts like a meter (see Meck and Church, 1983). After each counting event, the analogical representation grows in size and as the size grows, so does the estimation of the number size. Using an analogical representation, animals (e.g., Gallistel & Gelman, 1992) and children (e.g., Brannon & van de Walle, 2001) that cannot use verbal representations can still count.

An assumption of the above studies, particularly Experiment 2, is that participants use a verbal representation of number in order to count. The verbal working memory load condition was included in Experiment 2 to interfere with the representation of the current total. If the total was not represented in a verbal format, then there should be no interference or decline in counting performance.

Counting did decline when participants stored a verbal working memory load. This suggests that participants relied on a verbal representation of number and the load caused increased interference. This finding does not exclude the possibility that participants used both representations of number at the same time. If participants stored both, then they may have used the analog representation to approximate the current total after forgetting the verbal representation. However, this does not explain the trend towards undercounting in Experiment 2. There is no reason that the analog representation should provide a number that is too low as opposed to too high.
According to Cordes et al. (2001), people can use both types of representations at different times. When their participants counted under conditions suppressing their ability to store phonological information, they showed error patterns characteristic of analog number representation. When participants could store phonological information verbally, they elected to do so, showing error patterns common to verbal representation of number.

In Experiment 1, participants showed variability patterns similar to those of verbal counting in the Cordes et al. (2001) paper in all conditions but the 400 ms SOA condition. Even in the 400 ms condition, the pattern of variability was not conclusively analog counting. The variability patterns indicate that participants used verbal numbers as their primary number representation.

*ACT-R.* ACT-R (Anderson & Lebiere, 1998) is another important consideration of the current study because ACT-R affords tests and justifications for the theoretical assumptions made in the present research. ACT-R is a cognitive modeling system that attempts to represent all of human cognition under one set of principles and constructs. Researchers use ACT-R in computer simulations to approximate empirical data and justify their own more specific models. This approach is applicable to the current research given the parallel structure (outlined in the Introduction) between ACT-R and the counting model adopted in the current study (proposed in Carlson & Cassenti, in press).

The first step for a model of the present research in ACT-R is to model the findings of Carlson and Cassenti (in press), the first research to adopt the counting model used in this study. Carlson and Cassenti (in press) found that undercounting is the most
frequent type of error for relatively rapid rhythmic presentation of events (rhythmic presentation is overall less error-prone) and overcounting is the most frequent type of error for varied-interval presentation of events at similar rates. An ACT-R model that captures these findings includes a temporal estimation mechanism, activation noise that causes the current total to be selected as the prospective total on some steps (to explain undercounting), and a tendency to count again when in a varied-interval condition (to explain overcounting). A more detailed account of this first step to an ACT-R model is outlined in Appendix 1. To extend the model to the current study, new mechanisms must be added including the tendency to rehearse prospective totals during long intervals and infrequent distracters, which are sometimes counted in error.

Error Detection. The error monitoring data in Experiments 1 through 4 demonstrate that when a counting error is caused by a disruption to the flow of processing, participants are most likely to detect that an error has occurred. Causes of the disruptions included intention-outcome confusions during long SOA conditions (Experiment 1) and reconstruction of forgotten goal representations (Experiments 2 and 4). In Carlson and Cassenti (in press) detection of errors was greatest for varied-interval presentation of items, another disruptive condition.

Postma’s (2000) model of speech error suggests that error detection is triggered by disruption of the natural flow of sentence construction. Postma (2000) claimed that all sentence construction begins with an intention to construct a certain type of message and proceeds through a set of mental processes used to construct a proper sentence (i.e., a sentence that follows all rules and matches the intended message). When errors occur, language-monitoring processes detect them by detecting disruptions in the flow of mental
processes. Postma argues that disruptions in the flow of mental processing are caused by mismatches between intended message and constructed sentence or by inconsistencies between language production rules and constructed sentence. Instead of detecting errors explicitly by detecting the mismatch or inconsistency directly, monitoring processes detect an error through an effect of the mismatch or inconsistency, a disruption to flow of processing.

Postma’s (2000) model of speech errors matches the model of error detection in the present perspective. A goal representation to count is a simpler intention than the intention to produce a sentence with a given message, but both require that certain parameters be specified before enacting a procedure. For instance, the intended message must include at least a subject and verb and a counting intention must include the prospective total. The flow of processing in counting and in sentence production is disrupted when an intention is forgotten or when representations are not readily distinguished (e.g., confuse the prospective total for the current total or confuse two words and use the wrong one in a sentence).

Limitations of studies. One limitation of these studies was the match between Experiment 1 and Experiment 3. One of the potential reasons for overcounting at a long SOA was the random and infrequent appearance of distracters in the environment. In Experiment 1, participants overcounted in the long SOA, but there was no instruction to avoid counting distracters.

Experiment 3 was the follow-up to Experiment 1 and included distracters that appeared outside a box on the screen. Participants were instructed to count only items
that appeared inside the box. Including this instruction was a mismatch between Experiment 3 and Experiment 1, which did not have this instruction.

This difference between Experiment 1 and Experiment 3 created a difficult situation in interpreting the distracters hypothesis of overcounting because in Experiment 3 the distracter conditions did not show overcounting. Perhaps, the instruction to only count what appeared in the box caused participants to ignore the distracters. Experiment 4 showed that this may have been the case as participants who could not ignore the distracters tended to commit more overcount errors. Participants in Experiment 1 should have received a similar instruction to avoid counting distracters.

Another limitation was the low accuracy on the secondary pressing task in Experiment 4. In Experiment 2, I filtered out all trials in which participants did not remember all letters in the verbal load. A similar filter for all accurate pressing trials would have drastically reduced the power of all analyses.

Given this potential loss of power, no pressing filter applied to the data of Experiment 4. Including these trials also includes all trials in which participants were not motivated to perform well on the secondary task. A variation of Experiment 4 in which participants pressed less often or performed a less difficult task would have provided more convincing results.

Another limitation of Experiment 4 was the way in which press behavior was recorded. Instead of information about each individual press, the data reflected press behavior over the entire trial. One of the conjectures in the above analyses is that participants may have forgotten that they had pressed and press again to compensate for a perceived miss. With the current state of the data, there is no way to determine when
participants might have pressed the button a second time. For instance, it could be that participants pressed the button twice in quick succession. This finding would have indicated that participants were not monitoring whether or not they pressed, but just that they interacted with the mouse in an improper way.

*Future Directions.* There are several potential directions to follow up on this research. The theoretical concepts in this research need to be studied further given their importance to cognitive processing. For instance, the theory that the goal representation is a format of working memory distinct from the visual and verbal formats has far-reaching implications for the working memory model of Baddeley and Hitch (1974).

Recently, Baddeley and Logie (1999) indicated that the central executive should be the primary focus of working memory research. Goal representations implicate many directions for this research. The primary implication is that working memory is not a collection of distinct components with different functions, but different type of representations that do not interfere with one another. Therefore, a potential future direction is to frame studies of working memory in terms of the interference of different formats of working memory representations and to discover if goal representation loads interfere with central executive tasks.

Another potential future direction is a more in depth analysis of distracters. For instance, what types of distracters trigger intentions? Mixed results were found in Experiments 3 and 4, in which only distracters that were attended caused overcounts. Perhaps a more subtle manipulation of distracters would have caused more overcounts. For instance, participants could be instructed to count all items that appear on the screen, while a computer-controlled light flashes on an external device, near the monitor. This
manipulation would not require as strong an instruction as in Experiment 3 (i.e., to only count what appeared inside the box), but would still simulate the hypothesized infrequency of distracters. This experiment could be coupled with an experiment similar to Experiment 1, in which some participants are explicitly warned not to count any distracters. If warned participants make fewer overcounts than unwarned participants then, perhaps the lack of overcounting in Experiment 3 can be explained by actively ignoring the area in which distracters appear.

One other future direction is to investigate individual differences in working memory capacity more fully. A potential reason for lack of a significant trend of better counting performance with higher working memory capacity is that there was not enough power in terms of trials and subjects. The other more theoretical reason is that a test of verbal capacity does not accurately measure capacity for goal representations, the manipulation in Experiment 4.

More participants and more trials would help to distinguish which is the correct explanation. If an increase in number of observations does not reveal a significant trend, then a new measure of working memory capacity may be constructed, which measures working memory capacity for goal representations. Perhaps a goal representation measure would reveal a significant match between capacity and counting performance in a study like Experiment 4.

CONCLUSIONS

Several points from these studies stand out. First, more time to count actually hurt performance. This decline in performance was attributed to increased rehearsal of the
Increased rehearsal typically helps performance, but in the case of counting, rehearsal actually causes intention-outcome confusions, leading to errors.

The studies also revealed information about goal representations. Goal representations may be another format of working memory representation in addition to verbal and visual representations. This finding has implications for models of working memory. For instance, goal representations may be the special format guiding control of working memory.

Another finding about goal representations concerns how goal representations may be constructed and triggered. According to the above analysis, goal representations are built piecemeal. When the final piece of the goal representation becomes available, the goal representation becomes an intention and that intention triggers the operation specified by the intention. Individuals have some measure of control over when goal representations become intentions (e.g., they can ignore stimuli if instructed to), but when the focus of attention shifts to the item that can complete the goal representation, the operation triggers.

These findings imply other studies that will reveal more information about goal representations and the nature of cognitive control in working memory. This research is merely the starting point from which numerous studies may begin.

This study began by suggesting that the speed-accuracy tradeoff (Hale, 1969) was more than fast mental processing leading to lower accuracy. Indeed, the empirical studies demonstrated that slowing mental processing hurt performance, as did speeding up processing. This study has shown that fast or slow paces are not the correct way to
conceptualize the relation between processing time and error, but that non-optimal timing in either direction increases error.

Researchers have traditionally conceptualized the speed-accuracy tradeoff as a relation between speed of self-paced mental processing and performance. In this study, manipulation of event presentation altered the amount of time the environment provided for mental processing. Individuals constantly process events in the environment at paces not under their control and this study highlights the problems that can occur when the pace of events is not optimal. For example, consider a long-distance phone call in which speakers have a difficult time pacing their speech to accommodate listeners whose interactive feedback is delayed across distance. Clearly, errors are more common when communication transmission time makes adjusting pace of speech to a listener’s feedback more difficult.

Three hypotheses were proposed that explained why long processing time leads to lower accuracy. Forgetting was proposed as a factor because memory decay increases with more time. Second, over-rehearsal was thought to lead to more intention-outcome confusions. Lastly, distracters may lead to incorrect triggering of intentions. Of all the hypotheses, the data indicated that over-rehearsal was the most likely cause of error. When no working memory load reduced the chance to rehearse, the type of error characteristic of a long processing time (i.e., overcounts) appeared. This is not to say that the other factors did not influence counting. Clearly, forgetting and distracters do alter performance, when under conditions of working memory load (Experiments 2 and 4) and with focused attention of distracters (Experiment 4). I discuss implications of each of the three factors below.
The results show that rehearsing information that individuals use to enact their next intention can hurt performance of a cascaded task (i.e., a task requiring constant updating of information, such as in counting). Cascaded tasks are common in everyday life. For example, pilots must execute all the tasks in a checklist before landing (see Barshi & Healy, 1993). If a delay exists between one task and the next, pilots may rehearse the next task so many times that they may believe they had already achieved the objective and fail to execute a necessary task. Clearly, failure to execute a task in a landing procedure may have dire consequences.

Experiments 2 and 4 showed that working memory loads hurt performance when counting. Decrementsin performance were found with a verbal load of several representations and a goal representation load of just one extra goal representation. Individuals can expect drops in performance of even simple tasks like event counting by simultaneously storing extra information. People should be wary of enacting a secondary task (Experiment 4) and storing a prospective secondary task (Experiment 2) while enacting another task. Ability to avoid errors is best when focusing on only one task at a time, such as when a pilot is preparing to land.

Manipulations involving distracters also influenced the participants’ behavior. Distracters that capture attention can lead to incorrect triggering of intentions. This can lead to poor performance when an individual attempts to react to only certain types of information. As an example, consider a security officer who must respond to a crowd of protesters. If any loud noise triggers a violent response from the officer, a peaceful protest could turn into a riot. In this case, the officer must only respond aggressively
after a protester acts aggressively and illustrates the need to monitor which events are included in intentions.

Other important considerations of this study were the roles of goal representations and working memory in cognition in relation to the speed-accuracy tradeoff. The results showed that goal representations must be triggered with enough time to ensure that they are complete, yet soon enough that they are not susceptible to adverse effects from long processing time. Working memory is more susceptible to forgetting over long processing times that increase the chance for over-rehearsal (to counter-act increased decay) and interference from other goal representations, which are more likely to be activated over more time. Also, a long processing time increases the chance that a distracter will appear which erroneously triggers an intention stored in working memory. The role of goal representations and working memory are both important to understanding why a long processing time hurts performance.

Design of technology and design of training regimens are important applications of this research. People perform optimally when task-relevant events are presented with timing that matches their own mental processes or when individuals can control the timing of events in accordance with their own mental processes (e.g., Stevenson & Carlson, 2003). Applied cognitive work (e.g., Hembrooke & Gay, 2003) can use this research to improve people’s performance in many tasks by being mindful of optimal timing of mental processes, reducing goal representation and verbal working memory loads, and avoiding distracting information. In many ways, technology and training shape individuals’ lives and consideration of human cognition will help to improve those lives.
REFERENCES


APPENDIX A

ACT-R MODEL

An ACT-R (Adaptive Control of Thought, Anderson & Lebiere, 1998) model of event counting can contribute to the current research by justifying and outlining areas of improvement for the theoretical assumptions of the counting model (Carlson & Cassenti, in press). I will begin by describing the goal of modeling in general and for counting, then describe the data to be modeled, the components of the model, and how those components may be implemented in an ACT-R model.

Goal of Model

**General Purpose of Modeling.** Modeling cognition is a method for justifying a theoretical explanation for empirical results and allowing an objective standard for improving the explanation when it is erroneous. ACT-R and other cognitive modeling systems (such as Soar, Newell, 1990) provide frameworks for cognition that can be implemented in computer simulations. The computer simulations are derived from adjustments to the properties of the modeling system. For instance, in a counting model, a researcher would need to add to the model’s knowledge base by including knowledge of numbers.

The purpose of each computer simulation is to approximate the data from an empirical study. If the simulation is adjusted to the assumptions made by the researcher and the empirical results closely match the results of the simulations, then the modeling system justifies the theoretical assumptions.

**Purpose of Modeling Counting.** A model of simple counting (i.e., counting without attaching the numbers to specific items) has already been established (Anderson
& Lebiere, 1998), however event counting and specifically event counting with the assumptions made in the counting model (Carlson & Cassenti, in press) adopted here has yet to be established. An ACT-R event counting model is important to justifying the descriptive event counting model described in the Introduction.

Data to be Modeled

The starting point for a model of event counting is the empirical results of the first study to test the counting model by Carlson and Cassenti (in press). I will describe the important results for one of these experiments (Experiment 2, Carlson & Cassenti, in press) and explain a model of how these results may have occurred.

Rapid rhythmic and varied-interval presentation. The experiment to be modeled was much like Experiment 1 of the present study with differences only in timing of item presentation. Instead of four rhythmic SOAs, there were two types of rapid presentation rates. In a rhythmic condition, participants counted items presented at a pace of 550 ms between item onsets (350 ms of blank interval and 200 ms of asterisk item presentation). In the other condition, participants counted items presented with varied-interval timing, in which an interval from the set: 400, 500, 600, and 700 ms was randomly selected as the SOA on each step of a trial (note that the average duration was the same as the rhythmic interval).

At the end of each trial, participants typed the current total and their confidence that the typed total was correct. I will refer only to results involving the total response and not the confidence response.

Accuracy. As in Experiments 1 through 4 of the current study, trials were accurate if participants entered the exact, correct total of items. In the experiment to be
modeled, participants counted correctly on 83% of rhythmic trials and on 53% of varied-interval trials. These results reflect a definite advantage for rhythmic presentation of items.

Clearly, participants use some type of temporal mechanism to take advantage of the rhythmic structure of items and develop a temporal expectation for when a new event will appear. Perhaps participants use this expectation to time their own mental processes to coincide with the appearance of new items, providing optimal conditions for counting. This implication will be explored in the section on temporal components of model.

*Rhythmic undercounting and Varied Overcounting.* In addition to the difference in overall accuracy, Carlson and Cassenti (in press) found a difference between the conditions in type of error made. When participants made errors in the rhythmic condition, those errors tended to be undercounts. When participants made errors in the varied-interval condition, those errors tended to be overcounts. This dichotomy demonstrated a difference in cognitive processing when participants had a temporal expectation and when they did not.

Recall that there were fewer overall errors when items were presented rhythmically. If rhythmicity aids participants to count correctly, then perhaps undercounting is the more natural type of error and overcounting results from lacking a temporal expectation. This approach will be elaborated in the components of model section.

*ACT-R Modeling System*

Before considering the ACT-R event counting model, it is necessary to understand how ACT-R itself operates. The system of ACT-R is built on an architecture of cognitive
modules and processes that together form the core of any individual model in ACT-R. The main parts of ACT-R important here are declarative and procedural memory, the goal buffer, the retrieval buffer, and ACT-R/PM (PM stands for Perceptual-Motor).

ACT-R uses its modules in concert with one another. The starting point is the goal buffer, which contains the information ACT-R needs to select its next action. Production memory contains production rules, a list of individual actions that help ACT-R satisfy its goals. Production rules are if-then statements. The “if” part of a production rule specifies the conditions that must be met to enact the production and the “then” part specifies the actions that should take place once the conditions are satisfied. Declarative memory contains the pieces of information pertinent to the goals of ACT-R, called chunks. These chunks are requested by productions rules. Once a chunk is selected, the chunk is included in the retrieval buffer, which is a short-term store for retrieved chunks. Lastly, ACT-R/PM is the perceptual motor component of ACT-R. PM allows ACT-R to take information in from the environment (i.e., perception) and perform actions in the environment (i.e., motor).

Event counting provides a good illustration of the nature of each component in ACT-R. The goal buffer would contain the goal to count, the current total, and other information suggesting the next production rule that should fire (e.g., after seeing the item, the next step is to increment). The production that the goal indicates is the most likely one to be selected next. The actions of many productions are simply to request new information, either from declarative memory or from the perceptual component of ACT-R/PM. For instance, one production rule in the counting model could request that declarative memory select a total from memory and another production rule could request
that ACT-R/PM look for the item information. When the total or item information is retrieved, it is included in the retrieval buffer.

**Components of Model**

The data described in the previous section can be modeled with a set of three mechanisms. The temporal mechanism adds an explanation for how people develop a temporal expectation for a new item. The undercounting and overcounting mechanisms explain how this temporal mechanism results in the differences in types of error between timing conditions.

**Temporal mechanism.** Individuals must have a temporal expectation mechanism, which alters the quantity and type of error made in the Carlson and Cassenti (in press) data set between timing conditions. To my knowledge, this is the first attempt to include a temporal expectation mechanism in ACT-R. ACT-R can achieve temporal expectation by adjusting the time it takes to run a production rule. All production rules are initially set to take 50 milliseconds (ms), but cognitive modelers can increase this number to reflect when a production takes longer than others. If an event counting model is adjusted so that productions fire to process the event at the precise time that it appears, a temporal expectation mechanism will result.

Although ACT-R has the ability to achieve temporal expectation, it is also important to justify the conditions of the model theoretically, one of the primary functions of ACT-R. Baddeley, Thomson, and Buchanan (1975) found that the length of time to say a word is proportional to the amount of time to say the word in inner speech. This suggests that speaking a syllable both out loud and with inner speech takes a constant amount of time. To time counting so that it corresponds to the time it takes for
the next item to appear, ACT-R should have a set number of syllables to speak for each number that take the same length of time to speak as the next item takes to appear and incrementing the total just after speaking the old total.

There is reason to believe that people use mechanisms like inner speech to estimate time. Consider a musician playing an instrument, while tapping a foot. Jones and Boltz (1989) claim that the foot acts as an effector, a part of the body which aids time production. The theory claims that effectors are directly linked to an internal time keeper and aid the individual in anticipating when the next rhythmic event should take place.

I propose that another effector is inner speech. If syllables in inner speech take a constant amount of time to speak, then the individual can use inner speech to aid time production. This temporal expectation is the basis for aiding rhythmic counting and the misuse of this mechanism is the basis for why varied-interval timing hurts counting.

*Rhythmic counting*. Rhythmic counting is a less error-prone process because individuals have some type of temporal expectation mechanism that helps them avoid error. I propose that individuals speak a certain number of syllables for each number in inner speech (e.g., seventy-seven is sev-sev to abbreviate the number to two syllables, see Mueller, Seymour, Kieras, & Meyer, 2003) and once the numbers are spoken expect to see a new item and increment the current total. This mechanism allows them to prepare the new number in memory and make it the new count as soon as the item appears. A decay mechanism in ACT-R reduces the activation of declarative memories that linger too long in working memory. To avoid this decay, retrieval of the new number should occur just before the next item appears.
Even with the ability to increment in time with the appearance of the new item, the data form Carlson and Cassenti (in press) show that individuals do make errors and that those errors are generally undercounts. An ACT-R mechanism that can explain undercounts is activation noise. In ACT-R, a declarative chunk (e.g., a number) has an activation level as an indicator of the strength of the chunk and the probability that the chunk will be retrieved. Activation noise is random activation, which increases or decreases the strength of activation within a set range of magnitude. If two declarative chunks can be selected then the one with the highest activation including noise is selected by a retrieval command. Undercounting can happen when the noise parameter gives the current number more activation than the prospective number, even though decay has decreased the activation of the current number. This explanation gives a reason for the predominance of undercounts during rhythmic presentation of items, but does not explain why overcounting is predominant during varied-interval presentation.

*Varied-interval counting.* Overcounting is best explained by an additional property, which primarily affects varied-interval presentation of items. This approach suggests that the noise explanation of undercounting is still present in varied-interval conditions, but that a lack of rhythmicity contributes another problem that causes overcounts, which sometimes correct an undercount, but in the absence of undercount creates overcount errors.

With varied-interval presentation of items, a timing mechanism does not provide a reliable expectation because items could appear between 400 and 700 ms after the onset of the previous item. A lack of expectation puts the individual at risk for either retrieving the new number too early or too late. Retrieving a number too early is a preferable
problem to retrieving a number too late. If ACT-R retrieves a number too late, then there would be no number to provide for the new item and ACT-R would need to initiate number retrieval during a time when the next item could appear. Retrieving a number early only means that the number will need to be stored until it is ready to be used.

Clearly, retrieving a prospective number early is preferable to retrieving the number late and throughout the trial an individual will learn that there is a limit as to how early the item can appear. To optimally time the retrieval of a new item, participants and ACT-R should set the timing mechanism to retrieve the new number by 400 ms after the onset of the last item. A 400 ms expectation would coincide with the minimal time, so a new total would always be available for each new item.

Activating the new number early creates the overcounting characteristic of varied-interval timing. One of the production rules for event counting is to count every time an item appears. If the new total is retrieved too soon before the next item, then the participant may count again once the item actually appears. This situation would cause enough overcounts to overcome activation noise that lead to undercounts, but also cause overcounts for trials in which activation noise does not cause an undercount.

Declarative Memory

There are only two types of declarative chunks that ACT-R needs to count. Of course, ACT-R needs knowledge of numbers in order to count, but ACT-R also needs goals, one of the key properties of the discussion in the main text.

Numbers. In order to implement the model outlined above, ACT-R needs several pieces of information about each number used to count. ACT-R needs to know the name of the number and five pieces of syllable information. Recall that in order to explain the
Carlson and Cassenti (in press) results, ACT-R must have a timing mechanism that estimates both 550 ms for rhythmic-interval and 400 ms for varied-interval presentations. If temporal expectation proceeds from syllables, then a 550 ms expectation should have more syllables than a 400 ms expectation (note that production times are written into the code of a model and cannot be changed by ACT-R based on the timing conditions of a trial).

In the numbers of the event counting model, the number name is split into two syllables for the 400 ms expectation and also into three syllables for the 550 ms expectation. When production rules fire, one set will retrieve and speak two syllables for each number in the varied-interval trials and another set will retrieve and speak three syllables for each number in the rhythmic trials. In each number chunk, number names will be expanded (e.g., one becomes o-o-ne, for 550 ms), contracted (e.g., seventy-seven becomes sev-sev, for 400 ms), or not change (e.g., thirty becomes thir-ty for 400 ms) to accommodate the syllable constraints.

Goals. Goals as declarative chunks are much like the goal representations described in the discussion of the main text. Goals are used in the goal buffer of ACT-R and act to control the flow of productions and to store current step information. The main difference between the goal representations of ACT-R and the goal representations of the main text is that goals in ACT-R specify the current production rather than the current operator, “to assign.” For example, if ACT-R needed to say the first syllable of a number, the goal would specify “speak first syllable.” In Carlson and Cassenti’s (in press) descriptive counting model, speaking the first syllable would fit most naturally between plan and intention, indicating that counting was on the way to specifying a full
intention by speaking the first syllable. Using the temporal mechanism would not require a new goal representation in the descriptive counting model, but would require a modified goal representation in ACT-R (i.e., containing the same current total, but different production information).

In an ACT-R model of event counting, goals should have at least three slots: the current total, production information, and the timing conditions of the current trial. The current total slot specifies the number representing the label of the previous event and triggers activation for the next number in the sequence. The production information indicates which production should fire next. In the “if” portion of each production rule there is a condition, which needs to match the production information in the goal before the production can fire.

The timing condition information is a necessary portion of the goal. Without this information, ACT-R would not know what temporal expectation to use for counting on a given trial. This aspect of the goal is discussed further in the limitations of model section.

Production Rules

The production rules for an event counting model can be classified into three categories: productions controlling the temporal mechanism, productions involving perception, and productions controlling the flow of numbers. Each of these productions is crucial to controlling ACT-R’s counting process and to reproducing the results of Carlson and Cassenti (in press) experiment.

Temporal Mechanism. Recall that a temporal mechanism in ACT-R involves speaking syllables to increment in synchronization with the desired temporal expectation.
There are seven productions controlling speaking the numbers. The first two productions retrieve the syllables needed to speak the number, one for retrieving two syllables and one for retrieving three syllables. Two of the productions match the 400 ms temporal expectation and fire to speak the first syllable and the second syllable of numbers. The other three productions match the 550 ms expectation and fire to speak the first, second, and third syllables. The temporal mechanism begins after ACT-R increments the total. After the productions fire, ACT-R increments the current total.

Perceptual. Five perceptual productions are needed to process event information. To engage perception ACT-R needs to start by trying to find the item presented on the screen. Once the item is located an attend production needs to ensure that perception focuses on the information. The third and fourth productions encode the identity of the item. Once the item is encoded into memory, ACT-R can abandon perception and process the item. If the item is an asterisk, then the action part of one production sends the model to an increment production. If the item indicates the end of the sequence (the letter “f” will be displayed to call for a response), then the action part of the other production sends the model to a respond production.

The final perception production helps ACT-R by providing a wait production if no item is found in the display. When the wait production fires, it returns ACT-R back to the find-item production. Without this wait production, ACT-R would stop running the model because no other production would meet the conditions for firing.

Incrementing. There are two ways in which ACT-R increments the current total. First, when the temporal mechanism has spoken the syllables for the current total, ACT-R increments. ACT-R also increments after it has processed an item. This dual route to
incrementing is the mechanism resulting in overcounts in the 400 ms expectation (i.e., increment after 400 ms and increment again if the item appears at 700 ms instead).

Figure A1-1 presents a depiction of the flow of productions in an overcount step. In the rhythmic condition, the item appears just when the temporal mechanism finishes speaking the syllables, so ACT-R should only count once. Figure A1-2 presents a depiction of the flow of productions for a rhythmic step.

Figure A1. Sequence of production rules in a varied timing trial in which the expectation mechanism is set to 400 ms, but the item does not appear until 700 ms. The temporal mechanism calls the increment production before perception calls its own increment production after seeing the item, thus incrementing twice and resulting in an overcount.
Cognition: 

Perception: 

Figure A2. Sequence of production rules in a rhythmic timing trial. Perceptual and cognitive processes trigger the increment production at the same time, resulting in only one increment production per step.

The first production retrieves the number zero from the goal chunk to start each trial and does not fire again until the next trial. An increment production retrieves the next number and leads to the retrieve production of the temporal mechanism. One last production is the respond production, which instructs ACT-R to provide the current count as the final count if perception has encoded the call for a response.

*Model Limitations*

*Timing goal slot.* One of the portions of the event counting ACT-R model that is psychologically implausible is that the timing condition of the goal is included from the beginning of each trial. Presumably, individuals would need to process several events before discovering which temporal expectation to use for counting.

Boltz (1993) suggests that developing a temporal expectation takes time, but it is developed through perceptual mechanisms built into cognition. To the author’s knowledge ACT-R has not yet developed a perceptual mechanism to handle discovering
a temporal expectation. As such, the temporal pattern is not discovered by the ACT-R event counting model, but instead must be assumed as given.

Perceptual and cognitive processing. Another limitation of the event counting model is setting the temporal mechanism to co-occur with perception of the item. In ACT-R, productions may engage perception while engaging cognitive processes (called parallel processing, see Byrne & Anderson, 2003). Parallel processing means that ACT-R may run the temporal mechanism while perception locates and identifies an item in the visual display. However, in the event counting model, it is not obvious which perceptual functions must run at the same time as cognitive functions. The model will remain incomplete until the proper linking of perceptual and cognitive processing is achieved.

Model Conclusions

Work remaining. The event counting model is currently incomplete. Proper linking of cognitive and perceptual processing needs to be implemented before the model can be fully realized. At this point, undercounts occur with manipulations of decay and noise parameters, but only with simple counting (i.e., counting without assigning numbers to events) and not with event counting. In the event counting model, ACT-R can perceive events, but always counts twice when using the dual route to incrementing (temporal mechanism and seeing the event) because cognitive and perceptual processing are activated in separate productions. More work needs to be done to simultaneously engage perception and cognition and then to adjust the parameters of the model to match the rates of undercounting and overcounting found in Cassenti and Carlson (in press).

Model’s contribution to ACT-R. The main contribution of the event counting model is the ability of ACT-R to temporally expect events. This temporal mechanism
avoids structurally or architecturally altering ACT-R, but works by using production rules and speaking, attributes that ACT-R already possesses. As mentioned above, the ability to temporally expect has yet to be developed in an ACT-R model

**Future directions.** The event counting ACT-R model of the Carlson and Cassenti (in press) data is only the first step in a wider research project. The goal of the model described in this appendix is to lead to a model of the results presented in the main text. To achieve a model of the results of the present studies, new mechanisms need to be added to the current model. Among these new mechanisms is a rehearsal function, which adds too much activation to prospective totals leading to overcounts and a distracter component, which adds distracters to the perceptual display, triggering extra increment productions.

Another future direction is to include a mechanism, which develops a temporal expectation only after several items are displayed. This extra development would provide psychological plausibility and justification for the Carlson and Cassenti (in press) counting model.
APPENDIX B

DISTRIBUTION OF ERRORS GREATER THAN FIVE

The following are depictions for all experiments of the distributions of error greater than five in magnitude. The tables represent the frequency of errors for each experimental condition and the figures represent the distributions of quantity of error across experimental conditions.

Table B1.

Experiment 1, Frequency of absolute errors greater than five for each condition.

<table>
<thead>
<tr>
<th>SOA</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>400 ms</td>
<td>13</td>
</tr>
<tr>
<td>800 ms</td>
<td>10</td>
</tr>
<tr>
<td>1600 ms</td>
<td>5</td>
</tr>
<tr>
<td>3200 ms</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
</tr>
</tbody>
</table>

Figure B1. Frequency of absolute errors greater across conditions for errors of 6 through 10 and greater than 10 in Experiment 1.
Table B2.

Experiment 2, Frequency of absolute errors greater than five for each condition.

<table>
<thead>
<tr>
<th>Response Order</th>
<th>Letter Set</th>
<th>SOA</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocked</td>
<td>Different</td>
<td>800 ms</td>
<td>3</td>
</tr>
<tr>
<td>Blocked</td>
<td>Different</td>
<td>3200 ms</td>
<td>5</td>
</tr>
<tr>
<td>Blocked</td>
<td>Same</td>
<td>800 ms</td>
<td>4</td>
</tr>
<tr>
<td>Blocked</td>
<td>Same</td>
<td>3200 ms</td>
<td>5</td>
</tr>
<tr>
<td>Mixed</td>
<td>Different</td>
<td>800 ms</td>
<td>1</td>
</tr>
<tr>
<td>Mixed</td>
<td>Different</td>
<td>3200 ms</td>
<td>0</td>
</tr>
<tr>
<td>Mixed</td>
<td>Same</td>
<td>800 ms</td>
<td>0</td>
</tr>
<tr>
<td>Mixed</td>
<td>Same</td>
<td>3200 ms</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>18</td>
</tr>
</tbody>
</table>

Figure B2. Frequency of absolute errors greater across conditions for errors of 6 through 10 and greater than 10 in Experiment 2.
Table B3.

Experiment 3, Frequency of absolute errors greater than five for each condition.

<table>
<thead>
<tr>
<th>Distracter Probability</th>
<th>Item Type</th>
<th>SOA</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target-Based</td>
<td>Distinctive</td>
<td>800 ms</td>
<td>0</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Distinctive</td>
<td>3200 ms</td>
<td>2</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Identical</td>
<td>800 ms</td>
<td>3</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Identical</td>
<td>3200 ms</td>
<td>3</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Distinctive</td>
<td>800 ms</td>
<td>3</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Distinctive</td>
<td>3200 ms</td>
<td>1</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Identical</td>
<td>800 ms</td>
<td>2</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Identical</td>
<td>3200 ms</td>
<td>1</td>
</tr>
<tr>
<td>Zero</td>
<td>Distinctive</td>
<td>800 ms</td>
<td>2</td>
</tr>
<tr>
<td>Zero</td>
<td>Distinctive</td>
<td>3200 ms</td>
<td>1</td>
</tr>
<tr>
<td>Zero</td>
<td>Identical</td>
<td>800 ms</td>
<td>1</td>
</tr>
<tr>
<td>Zero</td>
<td>Identical</td>
<td>3200 ms</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>21</td>
</tr>
</tbody>
</table>

Figure B3. Frequency of absolute errors greater across conditions for errors of 6 through 10 and greater than 10 in Experiment 3.
Table B4.

Experiment 4, Frequency of absolute errors greater than five for each condition.

<table>
<thead>
<tr>
<th>Secondary Task</th>
<th>SOA</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>800 ms</td>
<td>1</td>
</tr>
<tr>
<td>None</td>
<td>3200 ms</td>
<td>1</td>
</tr>
<tr>
<td>Target-Press</td>
<td>800 ms</td>
<td>5</td>
</tr>
<tr>
<td>Target-Press</td>
<td>3200 ms</td>
<td>2</td>
</tr>
<tr>
<td>Distracter-Press</td>
<td>800 ms</td>
<td>13</td>
</tr>
<tr>
<td>Distracter-Press</td>
<td>3200 ms</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>25</td>
</tr>
</tbody>
</table>

Figure B4. Frequency of absolute errors greater across conditions for errors of 6 through 10 and greater than 10 in Experiment 4.

The frequency of size of errors were large at errors of 10 and greater than 10 suggesting that, many of these errors resulted from storing and forgetting the tens digit, unmotivated counting, or typing errors. There were no noteworthy trends in the frequency of error per
condition for any experiment, except for the large number of errors for Distracter-Press, 800 ms in Experiment 4. Given that many more errors of a magnitude of five or less occurred in this condition than all other conditions, this result is not overly surprising.
APPENDIX C

TABLES OF ALL STATISTICAL TESTS

The following tables represent all statistical tests performed in Experiments 1 to 4.

Table C1.

Experiment 1, Proportion correct ANOVA table (within-subjects factor).

<table>
<thead>
<tr>
<th>Factor</th>
<th>Sum of Squares (SS)</th>
<th>Degrees of Freedom (DF)</th>
<th>Mean Square (MS)</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>1.222</td>
<td>3</td>
<td>0.407</td>
<td>11.080</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Error</td>
<td>2.536</td>
<td>69</td>
<td>0.037</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C2.

Experiment 1, Proportion correct t-comparisons (within-subjects factor).

<table>
<thead>
<tr>
<th>Paired Difference</th>
<th>Mean</th>
<th>Standard Error of the Mean</th>
<th>t</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>400 ms – 3200 ms</td>
<td>-0.18417</td>
<td>0.075075</td>
<td>-2.45</td>
<td>23</td>
<td>0.022</td>
</tr>
<tr>
<td>800 ms – 3200 ms</td>
<td>0.095833</td>
<td>0.04725</td>
<td>2.03</td>
<td>23</td>
<td>0.054</td>
</tr>
<tr>
<td>1600 ms – 3200 ms</td>
<td>0.088333</td>
<td>0.044088</td>
<td>2.00</td>
<td>23</td>
<td>0.057</td>
</tr>
</tbody>
</table>

Table C3.

Experiment 1, Curve Fit analysis with SOA as the independent variable and proportion correct as the dependent variable.

<table>
<thead>
<tr>
<th>Power</th>
<th>Residual MS</th>
<th>DF</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>0.014</td>
<td>94</td>
<td>1.35</td>
<td>0.249</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.110</td>
<td>93</td>
<td>5.77</td>
<td>0.004</td>
</tr>
</tbody>
</table>
Table C4.

Experiment 1, Chi-Square test comparing undercounts and overcounts.

<table>
<thead>
<tr>
<th>SOA</th>
<th>Observed Undercounts (&lt;i&gt;U_{obs}&lt;/i&gt;)</th>
<th>Observed Overcounts (&lt;i&gt;O_{obs}&lt;/i&gt;)</th>
<th>Total</th>
<th>Expected Undercounts (&lt;i&gt;U_E&lt;/i&gt;)</th>
<th>Expected Overcounts (&lt;i&gt;O_E&lt;/i&gt;)</th>
<th>&lt;i&gt;χ^2&lt;/i&gt;</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>400 ms</td>
<td>27</td>
<td>19</td>
<td>46</td>
<td>25.30</td>
<td>20.70</td>
<td>0.254</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>800 ms</td>
<td>10</td>
<td>7</td>
<td>17</td>
<td>9.35</td>
<td>7.65</td>
<td>0.100</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>1600 ms</td>
<td>14</td>
<td>5</td>
<td>19</td>
<td>10.45</td>
<td>8.55</td>
<td>2.680</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>3200 ms</td>
<td>9</td>
<td>18</td>
<td>27</td>
<td>14.85</td>
<td>12.15</td>
<td>5.121</td>
<td>1</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>49</td>
<td>109</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C5.

Experiment 2, Proportion correct t-comparisons (within-subject factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>0.236</td>
<td>1</td>
<td>0.236</td>
<td>5.954</td>
<td>0.018</td>
</tr>
<tr>
<td>SOA * Response Order</td>
<td>0.009</td>
<td>1</td>
<td>0.009</td>
<td>0.223</td>
<td>0.638</td>
</tr>
<tr>
<td>SOA * Letter Set</td>
<td>0.025</td>
<td>1</td>
<td>0.025</td>
<td>0.631</td>
<td>0.430</td>
</tr>
<tr>
<td>SOA * Response Order * Letter Set</td>
<td>0.113</td>
<td>1</td>
<td>0.113</td>
<td>2.866</td>
<td>0.096</td>
</tr>
<tr>
<td>Error</td>
<td>2.256</td>
<td>57</td>
<td>0.040</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C6.

Experiment 2, Proportion correct ANOVA table (between-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>80.094</td>
<td>1</td>
<td>80.094</td>
<td>1020.782</td>
<td>0.000</td>
</tr>
<tr>
<td>Response Order</td>
<td>0.892</td>
<td>1</td>
<td>0.892</td>
<td>11.373</td>
<td>0.001</td>
</tr>
<tr>
<td>Letter Set</td>
<td>0.365</td>
<td>1</td>
<td>0.365</td>
<td>4.649</td>
<td>0.035</td>
</tr>
<tr>
<td>Response Order * Letter Set</td>
<td>0.189</td>
<td>1</td>
<td>0.189</td>
<td>2.407</td>
<td>0.126</td>
</tr>
<tr>
<td>Error</td>
<td>4.472</td>
<td>57</td>
<td>0.078</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table C7.

Experiment 2, Count response time ANOVA table (within-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>0.007</td>
<td>1</td>
<td>0.007</td>
<td>0.013</td>
<td>0.911</td>
</tr>
<tr>
<td>SOA * Response Order</td>
<td>0.002</td>
<td>1</td>
<td>0.002</td>
<td>0.004</td>
<td>0.950</td>
</tr>
<tr>
<td>SOA * Letter Set</td>
<td>0.134</td>
<td>1</td>
<td>0.134</td>
<td>0.255</td>
<td>0.616</td>
</tr>
<tr>
<td>SOA * Response Order * Letter Set</td>
<td>0.653</td>
<td>1</td>
<td>0.653</td>
<td>1.245</td>
<td>0.269</td>
</tr>
<tr>
<td>Error</td>
<td>29.889</td>
<td>57</td>
<td>0.524</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C8.

Experiment 2, Count response time ANOVA table (between-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>834.699</td>
<td>1</td>
<td>834.699</td>
<td>628.984</td>
<td>0.000</td>
</tr>
<tr>
<td>Response Order</td>
<td>6.308</td>
<td>1</td>
<td>6.308</td>
<td>4.754</td>
<td>0.033</td>
</tr>
<tr>
<td>Letter Set</td>
<td>2.442</td>
<td>1</td>
<td>2.442</td>
<td>1.840</td>
<td>0.180</td>
</tr>
<tr>
<td>Response Order * Letter Set</td>
<td>0.021</td>
<td>1</td>
<td>0.021</td>
<td>0.016</td>
<td>0.901</td>
</tr>
<tr>
<td>Error</td>
<td>75.642</td>
<td>57</td>
<td>1.327</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C9.

Experiment 2, Chi-Square test comparing undercounts and overcounts.

<table>
<thead>
<tr>
<th>Response Order</th>
<th>Letter Set</th>
<th>SOA</th>
<th>U_{Obs}</th>
<th>O_{Obs}</th>
<th>Total</th>
<th>U_{Exp}</th>
<th>O_{Exp}</th>
<th>χ^2</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocked</td>
<td>Different</td>
<td>800 ms</td>
<td>10</td>
<td>4</td>
<td>14</td>
<td>8.89</td>
<td>5.11</td>
<td>0.377</td>
<td>1</td>
<td>&gt;.05</td>
</tr>
<tr>
<td>Blocked</td>
<td>Different</td>
<td>3200 ms</td>
<td>11</td>
<td>10</td>
<td>21</td>
<td>13.34</td>
<td>7.66</td>
<td>1.127</td>
<td>1</td>
<td>&gt;.05</td>
</tr>
<tr>
<td>Blocked</td>
<td>Same</td>
<td>800 ms</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>3.81</td>
<td>2.19</td>
<td>0.474</td>
<td>1</td>
<td>&gt;.05</td>
</tr>
<tr>
<td>Blocked</td>
<td>Same</td>
<td>3200 ms</td>
<td>11</td>
<td>7</td>
<td>18</td>
<td>11.44</td>
<td>6.56</td>
<td>0.045</td>
<td>1</td>
<td>&gt;.05</td>
</tr>
<tr>
<td>Mixed</td>
<td>Different</td>
<td>800 ms</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1.27</td>
<td>0.73</td>
<td>0.158</td>
<td>1</td>
<td>&gt;.05</td>
</tr>
<tr>
<td>Mixed</td>
<td>Different</td>
<td>3200 ms</td>
<td>8</td>
<td>3</td>
<td>11</td>
<td>6.99</td>
<td>4.01</td>
<td>0.402</td>
<td>1</td>
<td>&gt;.05</td>
</tr>
<tr>
<td>Mixed</td>
<td>Same</td>
<td>800 ms</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>4.45</td>
<td>2.55</td>
<td>0.189</td>
<td>1</td>
<td>&gt;.05</td>
</tr>
<tr>
<td>Mixed</td>
<td>Same</td>
<td>3200 ms</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>3.81</td>
<td>2.19</td>
<td>1.016</td>
<td>1</td>
<td>&gt;.05</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>54</td>
<td>31</td>
<td>85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table C10.

Experiment 2, Letter-Typing time ANOVA table (within-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>12.246</td>
<td>1</td>
<td>12.246</td>
<td>5.303</td>
<td>0.025</td>
</tr>
<tr>
<td>SOA * Response Order</td>
<td>0.184</td>
<td>1</td>
<td>0.184</td>
<td>0.080</td>
<td>0.779</td>
</tr>
<tr>
<td>SOA * Letter Set</td>
<td>29.760</td>
<td>1</td>
<td>29.760</td>
<td>12.888</td>
<td>0.001</td>
</tr>
<tr>
<td>SOA * Response Order * Letter Set</td>
<td>0.944</td>
<td>1</td>
<td>0.944</td>
<td>0.409</td>
<td>0.525</td>
</tr>
<tr>
<td>Error</td>
<td>131.620</td>
<td>57</td>
<td>2.309</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C11.

Experiment 2, Letter-Typing time ANOVA table (between-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5256.617</td>
<td>1</td>
<td>5256.617</td>
<td>337.670</td>
<td>0.000</td>
</tr>
<tr>
<td>Response Order</td>
<td>24.933</td>
<td>1</td>
<td>24.933</td>
<td>1.602</td>
<td>0.211</td>
</tr>
<tr>
<td>Letter Set</td>
<td>393.331</td>
<td>1</td>
<td>393.331</td>
<td>25.266</td>
<td>0.000</td>
</tr>
<tr>
<td>Response Order * Letter Set</td>
<td>0.056</td>
<td>1</td>
<td>0.056</td>
<td>0.004</td>
<td>0.952</td>
</tr>
<tr>
<td>Error</td>
<td>887.338</td>
<td>57</td>
<td>15.567</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C12.

Experiment 3, Proportion correct ANOVA table (within-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>0.234</td>
<td>1</td>
<td>0.234</td>
<td>20.378</td>
<td>0.000</td>
</tr>
<tr>
<td>SOA * Distracter Probability</td>
<td>0.053</td>
<td>2</td>
<td>0.026</td>
<td>2.287</td>
<td>0.108</td>
</tr>
<tr>
<td>SOA * Item Type</td>
<td>0.006</td>
<td>1</td>
<td>0.006</td>
<td>0.513</td>
<td>0.476</td>
</tr>
<tr>
<td>SOA * Distracter Probability * Item Type</td>
<td>0.026</td>
<td>2</td>
<td>0.013</td>
<td>1.118</td>
<td>0.332</td>
</tr>
<tr>
<td>Error</td>
<td>0.965</td>
<td>84</td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C13.

Experiment 3, Proportion correct ANOVA table (between-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>138.601</td>
<td>1</td>
<td>138.601</td>
<td>4361.879</td>
<td>0.000</td>
</tr>
<tr>
<td>Distracter Probability</td>
<td>0.163</td>
<td>2</td>
<td>0.082</td>
<td>2.567</td>
<td>0.083</td>
</tr>
<tr>
<td>Item Type</td>
<td>0.082</td>
<td>1</td>
<td>0.082</td>
<td>2.592</td>
<td>0.111</td>
</tr>
<tr>
<td>Distracter Probability * Item Type</td>
<td>0.058</td>
<td>2</td>
<td>0.029</td>
<td>0.918</td>
<td>0.403</td>
</tr>
<tr>
<td>Error</td>
<td>2.669</td>
<td>84</td>
<td>0.032</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

126
### Table C14.

Experiment 3, Signed error ANOVA table (within-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>0.370</td>
<td>1</td>
<td>0.370</td>
<td>7.881</td>
<td>0.006</td>
</tr>
<tr>
<td>SOA * Distracter Probability</td>
<td>0.012</td>
<td>2</td>
<td>0.006</td>
<td>0.128</td>
<td>0.880</td>
</tr>
<tr>
<td>SOA * Item Type</td>
<td>0.043</td>
<td>1</td>
<td>0.043</td>
<td>0.915</td>
<td>0.342</td>
</tr>
<tr>
<td>SOA * Distracter Probability * Item Type</td>
<td>0.015</td>
<td>2</td>
<td>0.007</td>
<td>0.157</td>
<td>0.855</td>
</tr>
<tr>
<td>Error</td>
<td>3.943</td>
<td>84</td>
<td>0.047</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table C15.

Experiment 3, Signed error ANOVA table (between-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.069</td>
<td>1</td>
<td>0.069</td>
<td>1.178</td>
<td>0.281</td>
</tr>
<tr>
<td>Distracter Probability</td>
<td>0.079</td>
<td>2</td>
<td>0.040</td>
<td>0.679</td>
<td>0.510</td>
</tr>
<tr>
<td>Item Type</td>
<td>0.093</td>
<td>1</td>
<td>0.093</td>
<td>1.598</td>
<td>0.210</td>
</tr>
<tr>
<td>Distracter Probability * Item Type</td>
<td>0.066</td>
<td>2</td>
<td>0.033</td>
<td>0.561</td>
<td>0.573</td>
</tr>
<tr>
<td>Error</td>
<td>4.910</td>
<td>84</td>
<td>0.058</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table C16.

Experiment 3, Chi-Square test comparing undercounts and overcounts.

<table>
<thead>
<tr>
<th>Distracter Probability</th>
<th>Item Type</th>
<th>SOA</th>
<th>U_{Obs}</th>
<th>O_{Obs}</th>
<th>Total</th>
<th>U_{Exp}</th>
<th>O_{Exp}</th>
<th>\chi^2</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target-Based</td>
<td>Distinctive</td>
<td>800 ms</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>4.19</td>
<td>2.81</td>
<td>0.393</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Distinctive</td>
<td>3200 ms</td>
<td>5</td>
<td>12</td>
<td>17</td>
<td>10.17</td>
<td>6.83</td>
<td>3.593</td>
<td>1</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Identical</td>
<td>800 ms</td>
<td>14</td>
<td>3</td>
<td>17</td>
<td>10.17</td>
<td>6.83</td>
<td>3.593</td>
<td>1</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Target-Based</td>
<td>Identical</td>
<td>3200 ms</td>
<td>15</td>
<td>19</td>
<td>34</td>
<td>20.34</td>
<td>13.66</td>
<td>3.485</td>
<td>1</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Distinctive</td>
<td>800 ms</td>
<td>7</td>
<td>5</td>
<td>12</td>
<td>7.18</td>
<td>4.82</td>
<td>0.011</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Distinctive</td>
<td>3200 ms</td>
<td>16</td>
<td>14</td>
<td>30</td>
<td>17.94</td>
<td>12.06</td>
<td>0.524</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Identical</td>
<td>800 ms</td>
<td>17</td>
<td>2</td>
<td>19</td>
<td>11.36</td>
<td>7.64</td>
<td>6.954</td>
<td>1</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Time-Based</td>
<td>Identical</td>
<td>3200 ms</td>
<td>15</td>
<td>15</td>
<td>30</td>
<td>17.94</td>
<td>12.06</td>
<td>1.202</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Zero</td>
<td>Distinctive</td>
<td>800 ms</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>4.19</td>
<td>2.81</td>
<td>1.954</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Zero</td>
<td>Distinctive</td>
<td>3200 ms</td>
<td>10</td>
<td>4</td>
<td>14</td>
<td>8.37</td>
<td>5.63</td>
<td>0.786</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Zero</td>
<td>Identical</td>
<td>800 ms</td>
<td>12</td>
<td>2</td>
<td>14</td>
<td>8.37</td>
<td>5.63</td>
<td>3.907</td>
<td>1</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Zero</td>
<td>Identical</td>
<td>3200 ms</td>
<td>6</td>
<td>7</td>
<td>13</td>
<td>7.78</td>
<td>5.22</td>
<td>1.009</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>128</td>
<td>86</td>
<td>214</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table C17.

Experiment 4, Proportion correct ANOVA table (within-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>0.003</td>
<td>1</td>
<td>0.003</td>
<td>0.071</td>
<td>0.791</td>
</tr>
<tr>
<td>SOA * Secondary Task</td>
<td>0.328</td>
<td>2</td>
<td>0.164</td>
<td>3.863</td>
<td>0.029</td>
</tr>
<tr>
<td>Error</td>
<td>1.781</td>
<td>42</td>
<td>0.042</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table C18.

Experiment 4, Proportion correct ANOVA table (between-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>57.725</td>
<td>1</td>
<td>57.725</td>
<td>1438.580</td>
<td>0.000</td>
</tr>
<tr>
<td>Secondary Task</td>
<td>1.163</td>
<td>2</td>
<td>0.581</td>
<td>14.488</td>
<td>0.000</td>
</tr>
<tr>
<td>Error</td>
<td>1.685</td>
<td>42</td>
<td>0.040</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table C19.

Experiment 4, Chi-Square test comparing undercounts and overcounts.

<table>
<thead>
<tr>
<th>Secondary Task</th>
<th>SOA</th>
<th>U_Obs</th>
<th>O_Obs</th>
<th>Total</th>
<th>U_Exp</th>
<th>O_Exp</th>
<th>$\chi^2$</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>800 ms</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1.18</td>
<td>1.82</td>
<td>0.045</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>None</td>
<td>3200 ms</td>
<td>12</td>
<td>11</td>
<td>23</td>
<td>9.04</td>
<td>13.96</td>
<td>1.602</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Target-Press</td>
<td>800 ms</td>
<td>2</td>
<td>16</td>
<td>18</td>
<td>7.07</td>
<td>10.93</td>
<td>5.990</td>
<td>1</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Target-Press</td>
<td>3200 ms</td>
<td>9</td>
<td>16</td>
<td>25</td>
<td>9.82</td>
<td>15.18</td>
<td>0.113</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Distracter-Press</td>
<td>800 ms</td>
<td>19</td>
<td>39</td>
<td>58</td>
<td>22.79</td>
<td>35.21</td>
<td>1.036</td>
<td>1</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Distracter-Press</td>
<td>3200 ms</td>
<td>23</td>
<td>18</td>
<td>41</td>
<td>16.11</td>
<td>24.89</td>
<td>4.858</td>
<td>1</td>
<td>&lt; .05</td>
</tr>
<tr>
<td>Total</td>
<td>66</td>
<td>102</td>
<td>168</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C20.

Experiment 4, Press accuracy ANOVA table (within-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>0.276</td>
<td>1</td>
<td>0.276</td>
<td>18.636</td>
<td>0.000</td>
</tr>
<tr>
<td>SOA * Secondary Task</td>
<td>0.015</td>
<td>1</td>
<td>0.015</td>
<td>0.999</td>
<td>0.326</td>
</tr>
<tr>
<td>Error</td>
<td>0.415</td>
<td>28</td>
<td>0.015</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table C21.

Experiment 4, Press accuracy ANOVA table (between-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>117.600</td>
<td>1</td>
<td>117.600</td>
<td>8.414</td>
<td>0.007</td>
</tr>
<tr>
<td>Secondary Task</td>
<td>30.817</td>
<td>1</td>
<td>30.817</td>
<td>2.205</td>
<td>0.149</td>
</tr>
<tr>
<td>Error</td>
<td>391.333</td>
<td>28</td>
<td>13.976</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C22.

Experiment 4, Median press absolute error ANOVA table (within-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>27.338</td>
<td>1</td>
<td>27.338</td>
<td>11.642</td>
<td>0.002</td>
</tr>
<tr>
<td>SOA * Secondary Task</td>
<td>16.538</td>
<td>1</td>
<td>16.538</td>
<td>7.043</td>
<td>0.013</td>
</tr>
<tr>
<td>Error</td>
<td>65.750</td>
<td>28</td>
<td>2.348</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C23.

Experiment 4, Median press absolute error ANOVA table (between-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>343.204</td>
<td>1</td>
<td>343.204</td>
<td>33.284</td>
<td>0.000</td>
</tr>
<tr>
<td>Secondary Task</td>
<td>5.704</td>
<td>1</td>
<td>5.704</td>
<td>0.553</td>
<td>0.463</td>
</tr>
<tr>
<td>Error</td>
<td>288.717</td>
<td>28</td>
<td>10.311</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table C24.

Experiment 4, Press signed error ANOVA table (within-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOA</td>
<td>43.350</td>
<td>1</td>
<td>43.350</td>
<td>12.535</td>
<td>0.001</td>
</tr>
<tr>
<td>SOA * Secondary Task</td>
<td>8.067</td>
<td>1</td>
<td>8.067</td>
<td>2.333</td>
<td>0.138</td>
</tr>
<tr>
<td>Error</td>
<td>96.833</td>
<td>28</td>
<td>3.458</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table C25.

Experiment 4, Press signed error ANOVA table (between-subjects factors).

<table>
<thead>
<tr>
<th>Factor</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>117.600</td>
<td>1</td>
<td>117.600</td>
<td>8.414</td>
<td>0.007</td>
</tr>
<tr>
<td>Secondary Task</td>
<td>30.817</td>
<td>1</td>
<td>30.817</td>
<td>2.205</td>
<td>0.149</td>
</tr>
<tr>
<td>Error</td>
<td>391.333</td>
<td>28</td>
<td>13.976</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX D

EXPERIMENT 2 RESULTS WITHOUT LOAD ERROR FILTER

The results section of Experiment 2 presented results after filter for all trials in which participants responded with a letter set that was not completely accurate. Give the size of this filter (19.2% of all data), the following is an account of the results before this filter with some discussion to compare the before and after filter results.

Accuracy. Only responses that were the same as the correct total were considered accurate. All other responses were considered errors. Participants made correct responses on 83.0% of all trials. Figure D1 shows the distribution of proportion correct by SOA, letter set, and response order conditions.

Figure D1. Proportion correct as a function of SOA and Letter Set-Response Order condition.

A mixed analysis of variance (ANOVA) on the within-subjects levels of SOA (800 and 3200 ms) and the between-subjects levels of letter set (different and same) and
response order (blocked and mixed) demonstrated a reliable difference in counting accuracy between SOA conditions, $F(1,60) = 22.92, MS_E = 0.70, p < .001$, letter set, $F(1,60) = 8.25, MS_E = 0.57, p = .006$, and response orders, $F(1,60) = 4.49, MS_E = 0.31, p = .038$. These significant main effects revealed that the 800 ms SOA condition (0.879) was more accurate than the 3200 ms condition (0.732); the same letter set (0.872) was more accurate than the different letter set (0.739); and mixed response order (0.855) was more accurate than blocked response order (0.756).

The ANOVA on proportion correct also revealed a significant interaction between letter set and SOA, $F(1,60) = 6.96, MS_E = 0.21, p = .011$, and a 3-way interaction, $F(1,60) = 6.96, MS_E = 0.21, p = .011$. Errors bars were calculated in accord with a method by Masson and Loftus (2003) and are shown in Figure 5, above. The difference between these results and after the load accuracy filter (as presented in the main text) was that the Blocked-Different-800 ms accuracy was greater than Blocked-Different-3200 ms and the Mixed-Different-3200 ms accuracy was greater than Blocked-Different-3200 ms. These two differences would not have had a profound effect on the interpretation of Experiment 2.

Response Time. A mixed ANOVA on all conditions was also run with count-typing response time as the dependent variable. The ANOVA revealed that the blocked response order (2390 ms) was faster than the mixed response order (2850 ms), $F(1,60) = 4.75, MS_E = 6.31, p = .033$. The ANOVA also revealed that the same letter set condition (2489 ms) was faster than the different letter set condition (2865 ms), $F(1,60) = 3.03, MS_E = 4.55, p = .087$. This letter set difference was not discovered after the load
accuracy filter, but would not have had a large effect on the interpretation of Experiment 2.

Signed Error. As in Experiment 1, signed error was calculated by subtracting the correct answer from the participant’s response. A chi-square test was conducted on the frequencies of undercount and overcount errors. Table D1 represents the frequencies of undercounts and overcounts in each condition.

Table D1.

Frequency of undercount and overcount errors in Experiment 2.

<table>
<thead>
<tr>
<th>Response Order</th>
<th>Letter Set</th>
<th>SOA</th>
<th>Undercount</th>
<th>Overcount</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocked</td>
<td>Different</td>
<td>800</td>
<td>12</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Blocked</td>
<td>Different</td>
<td>3200</td>
<td>16</td>
<td>13</td>
<td>29</td>
</tr>
<tr>
<td>Blocked</td>
<td>Same</td>
<td>800</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Blocked</td>
<td>Same</td>
<td>3200</td>
<td>12</td>
<td>7</td>
<td>19</td>
</tr>
<tr>
<td>Mixed</td>
<td>Different</td>
<td>800</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Mixed</td>
<td>Different</td>
<td>3200</td>
<td>14</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>Mixed</td>
<td>Same</td>
<td>800</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Mixed</td>
<td>Same</td>
<td>3200</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>70</td>
<td>42</td>
<td></td>
<td>112</td>
</tr>
</tbody>
</table>

Note. Only includes errors of five or less in magnitude.

All undercount and overcount errors were added across SOA conditions to determine the overall likelihood of each type of error. Overall, there was a 62.5 percent chance of an undercount and a 37.5 percent chance of an overcount. Each condition was tested individually to check if the frequency of undercounts and overcounts differed from this likelihood proportion. No conditions differed significantly from the overall proportion. This was the same finding after the load accuracy filter.

Error Monitoring. Confidence ratings on a scale of 1 to 5 were divided into low confidence (ratings of 1 to 3) and high confidence (ratings of 4 and 5) for each condition. A proportion of low and high confidence was calculated for both error and correct trials.
Figure D2 displays this proportion for error trials. Figure D3 presents this proportion for correct trials.

Figure D2. Proportion of error trials with low and high confidence as a function of SOA condition.
Figure D3. Proportion of correct trials with low and high confidence as a function of condition.

Figure D2 illustrates that when errors were made in the Blocked condition the participants showed a substantially greater chance of recognizing that the error was made, by rating the confidence in the count response as low. Figure D3 illustrates that when errors were not made, all conditions, except possibly the Blocked-Different conditions were rated with high confidence.

The biggest difference between these figures and the ones derived after the load accuracy filter was the increase in detected errors for the mixed-different-3200 ms SOA. This difference is not especially surprising given that the lack of a load accuracy filter includes trials in which participants forgot some portion of the letter set. Forgetting a
portion of the letter set may have been a large enough disruption to decrease reported
confidence for the count.

*Letter-Typing Time.* A mixed ANOVA was run on all conditions of Experiment 2
with letter-typing time as the dependent measure and revealed that typing the same letter
set (4959 ms) was faster than typing a different letter set (9580 ms), $F(1,60) = 43.83$, $MS_E$
$= 683.58, p < .001$. After the filter, there was also a difference in SOA and an interaction
between letter set and SOA not found in this test. These latter two findings were not
especially important to the discussion following these findings in the main text.
Vita
Daniel Nicholas Cassenti

Current Address:  Campus Address:
1605 Highlandon Ct.  604 Moore Building
State College, PA 16801  University Park, PA 16801

Phone number:  (814) 278-7777    (814) 863-9729

Email address: dnc112@psu.edu

Education
• 2004 Doctor of Philosophy Psychology at The Pennsylvania State University
• 2001 Master of Science Psychology at The Pennsylvania State University
• 1999 Bachelor of Science Psychology with Honors, Statistics Minor at University of Connecticut, Storrs, Magna Cum Laude

Publications

Presentations
• November 2003, Annual Meeting of the Psychonomic Society: poster presentation, *Temporal Dynamics of Deliberate Control*
• March 2003, Eastern Psychological Association: paper presentation, *Placekeeping in Counting Tasks*
• March 2002, Graduate Student Research Exhibition at Penn State University: poster presentation, *Counting and Place Keeping*
• April 1999, Cognitive Science Meeting at University of Connecticut: poster presentation, *Bowling for Words: A Dynamic Approach to Word Recognition*

Distinctions and Honors
• 2003 (Fall) Penn State Liberal Arts Travel Award
• 2003 (Spring) Penn State Liberal Arts Travel Award
• 2000 Graduate Certificate: ACT-R Summer School, Carnegie Mellon University
• 1999 Psi Chi National Psychology Honors Fraternity
• 1997 and 1998 New England Scholar Awards
• 1997 Awarded Sophomore Honors Certificate