

Typing and Talking: draft in progress

Matthew J. C. Crump<sup>1,2</sup>, Nicholas Brosowsky<sup>2</sup>, & Lawrence Behmer<sup>1</sup>

<sup>1</sup> Brooklyn College of the City University of New York

<sup>2</sup> Graduate Center of the City University of New York

#### Author Note

<https://github.com/CrumpLab/TalkTyping> is the github repository for this project containing the data, and source code for compiling compiling this paper in R using papaja.

Correspondence concerning this article should be addressed to Matthew J. C. Crump, Brooklyn College of CUNY, 2900 Bedford Avenue, Brooklyn, NY, 11210. E-mail: [mcrump@brooklyn.cuny.edu](mailto:mcrump@brooklyn.cuny.edu)

## Abstract

Enter abstract here. Each new line herein must be indented, like this line.

*Keywords:* Typing, Language production

Word count: X

## Typing and Talking: draft in progress

Language production processes turn ideas into actions. For example, people use utterances and gestures to convey semantic meaning during conversation; to preserve ideas in different mediums, such as hand-writing, typewriting, and dictation; and to plan and guide cognition and behavior in general. In Vygotsky's () developmental tradition, children first acquire connections between actions and outcomes in the world around them with the words spoken to them by their caregivers, and then internalize the same language system as a tool to regulate their own cognition and behavior. In adulthood, people continue to talk out loud and silently to themselves using inner speech, and such self-directed language production is assumed to continue to regulate many aspects of cognition and behavior (). For example, in memory, the phonological loop transfers to-be-remembered information into long-term memory through short-term rehearsal (Baddeley & Hitch, 1974). The act of speaking out loud itself improves future retrieval, as shown by the production effect (MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010); and, the process of naming may guide category learning by drawing attention toward diagnostic features (Lupyan,). Verbal recoding can facilitate learning and recall of lengthy sequences by casting elements into chunks with recallable names (Miller, 1956). In the completion of everyday tasks, Botvinick and Plaut (2004) argue that verbally coded instructions together with recurrent associative learning of real-world action sequences provide the control necessary for accomplishing routine tasks like making coffee or tea. Among these and other uses, a crowning achievement of language production is the ability to control the serial order of utterances to convey meaning in highly complex and flexible ways. In the present work, we are interested how the verbal process of language production for speech may or may not be recruited to facilitate manual control of fingers for sequencing keystrokes during typewriting. In particular, we investigate the general idea that the sequence production processes enabling fluent typing are controlled by common word level linguistic codes used during speech.

We review three perspectives from the typing, language production, and sequence production literatures that all suggest an important role for verbal codes to mediate keystroke production in typing. Typing is assumed to be controlled by hierarchically nested control loops (Logan & Crump, 2011). The outer loop relies on language comprehension and production to turn ideas into words and sentences. Word level-representation are then sent as goals to an inner loop that controls individual finger movements to individual keystrokes for each letter in a word. The two loops are informationally encapsulated modules (Fodor, 1983), and the outer loop does not know the details of how the inner loop accomplishes keystrokes (Liu, Crump, & Logan, 2010; Logan & Crump, 2009). However, the two loops must nevertheless be coordinated to accomplish accurate typing, presumably by intermediate word-level representations. For example, in Rumelhart and Norman (1982) typing model, words cause parallel activation of letter units in a response scheduling buffer, which then controls individual keystrokes for all letters in the word. That model focused on copy-typing, and assumed that word representations are activated by word-recognition processes operating during reading of to-be-copied text. In composition typing, word representations could be activated by overt or covert language production processes, and their activation would presumably have the same consequence of activating constituent letters for typing.

People are capable of overt and covert speech, which may be mediated by a common language production process. For example, according to the flexible abstraction hypothesis (), the selection of individual phonemes during covert or overt speech is controlled by a common process with relatively abstract phonemic codes. These abstract codes retain deep semantic and lexical features, but lack many of the articulatory features necessary for speaking out loud. As a result, inner speech may rely on these more abstract codes, which explains evidence that errors in overt speech can be driven more by lexical than phonemic similarity (). At the same time, these abstract codes can be used flexibly and further modified by articulatory processes that fill in necessary additional features to achieve overt speech. Although the flexible abstraction hypothesis is concerned mainly with speech overt

and covert speech production, it is worth considering whether it extends to nonverbal language production such as typing. For example, the same abstract verbal codes mediating inner speech may be co-opted and further specified by motor processes controlling finger movements to achieve accurate and fluid keystroke production. Sequence production more generally may also be mediated by verbal codes that trigger a plan-based mode of action control distinct from a more automatic stimulus-based mode of control (). For example, a stimulus-based mode of control refers to actions triggered in an automatic, reflex-like manner in response to well-learned cues from the environment. Whereas, a plan-based mode of control refers to goal-driven cognitive processes that set and supervise motor programs for complex sequencing of actions. Evidence for a distinction between control modes comes from the serial reaction time task (SRT, Nissen & Bullemer, 1987). The major finding from SRT tasks is that subjects learn to respond faster to targets that are presented in repeating or probabilistic sequences compared to random sequences, even if they are unable to explicitly describe the sequential regularities. At the same time, whether reaction times are sensitive to the frequency of particular targets or first-order transitions (e.g., repetitions or alternations) appears to depend on whether subjects possess explicit knowledge about sequential regularities in the task. For example, in a task where alternations were three times more likely than repetitions, subjects who were not aware of the manipulation showed faster responses to the more frequent alternations than repetitions, whereas subjects who were aware of the manipulation were equally fast for alternations and repetitions. One interpretation is that explicit knowledge of the transition probabilities allowed subjects to adopt a control-based mode, enabling action plans to mediate responses rather than more automatic frequency sensitive learning processes. Using a similar task, Tubau, Hommel, & López-Moliner (2007) further showed that subjects were more likely to adopt a plan-based mode (as evidenced by the elimination of the alternation benefit) when visual/verbal targets were used and when responses were followed by sounds, and plan-based control was disrupted when auditory noise was introduced. They suggested that use of phonemic codes

in general may strongly influence subjects to adopt a plan-based vs. stimulus-based mode of action control.

Several lines of evidence suggest that verbal codes for speech production participate in manual control over keystrokes during typing. Typing speed is fastest for words and decreases as letter strings approach random texts (). This effect of string structure could simply reflect familiarity and prior experience with typing words compared to random letter sequences, but could also be driven by string pronounceability. For example, typing speed is also influenced by syllabic structure, with faster interkeystroke times for letters within versus between syllables in a word. Word-level representations may also have a special status in activating all constituent letters in parallel compared to non-words which do not activate all letters in a sequence (). Finally, lexical structure at the level of bigrams appears to influence the spatial precision with which typists locate keys during typing (). This evidence is consistent with the idea that typists use verbal codes in a straightforward manner during typing, by concurrently using their inner voice (and perhaps sometimes speaking out loud) to say the same words they are typing.

Other work looking at dual-task interference in typing suggests that fast and accurate typing may not necessarily depend on using verbal codes. Protection from dual-task interference is often interpreted as evidence that a skill is highly automatized, and typing in general appears to be highly automatized. For example, typing speed in a dual task-condition to identify auditory tones with a foot-pedal was only 4ms slower than normal typing. Remarkably, Shaffer (1975) showed that typing speed was mostly unaffected under concurrent task demands to recite an unrelated nursery rhyme while typing. However, typing speed was slowed substantially when typists copied spoken words while simultaneously reading unrelated visually presented text. And, these data were collected only from a single skilled typist, and whether the result is generalizable remains unclear. On the one hand, if the same verbal codes used for speech are used for typing, it is surprising that a typist could

easily type one set of words while saying another set of words out loud without interference. On the other hand, interference appears to depend on the details of how verbal codes are used, as interference was observed when typists attempted to read unrelated text while copying spoken words.

The purpose of the present work is to clarify whether and how typists use verbal codes during typing. One possibility is that verbal codes for speech do not participate directly in the control of keystrokes for typing. On this view, typing could be considered a highly automatized and modularized skill that runs largely independently from other language production processes. Although speaking and typing may often go hand in hand, as many typists report silently voicing the words they are typing while they produce text (see Experiment 1), overt or covert speech production during typing could be carried out independently and in parallel with the motor movements required for typing. Another general possibility is that typists co-opt existing speech production processes during typing, so that normal typing is deeply connected at some level to a common language production process. We test these two general ideas in a series of related typing tasks that manipulate what typists say while they are typing. A central empirical aim was to tax concurrent speech production processes with reciting relevant or irrelevant content during copy-typing. In this way, we could establish boundary conditions for observing interference (or the lack thereof) from speech production on keystroke production.

In Experiment 1a and 1b we had typists speak the words or letters they were typing either silently or out loud. This experiment tested the idea that typists use verbal codes at the word rather than letter level during typing. The remaining experiments investigated the extent to which typing performance can be impaired by disrupting normal speech production. Experiment 2 measured copy-typing performance across several verbal suppression conditions where typists were instructed to repeat words or letter (unrelated to the text they were copying) out loud as they typed. Experiment 3 measured copy-typing performance

under conditions of delayed auditory feedback, which is known to disrupt fluent speech.

## Overview of present study

### Experiment 1a and 1b: Speaking words or letters during typing

The immediate goal of typing is to produce letters in the correct order during word and sentence production. Based on the idea that verbal codes are used during typing, we assumed that typists routinely verbalize, either covertly or overtly, aspects of the letters they are typing. For example, as I type this sentence, my subjective experience is that I silently voice the words I am typing. In Experiment 1a, we asked typists recruited online from Amazon's mechanical turk to copy-type a paragraph under two conditions to silently voice each word as they typed, or silently voice each letter as they typed. To gain an understanding of the subjective experience of using the inner voice during typing, we also asked typists to report what their inner voice usually says during typing (e.g., words, letters, both, or none). Because we had no way of controlling whether our subjects adopted our instructions, we conducted Experiment 1b as a lab-based replication and extension of Experiment 1a. Here, subjects copied five paragraphs. The first paragraph involved only the instruction to copy-type as quickly and accurately as possible. Across the remaining four paragraphs we manipulated whether subjects spoke out loud or silently to themselves, combined factorially with whether they spoke words or letters during typing. If typists use verbal codes at the word level during typing, then we expected that typing speed should be faster when subjects were instructed to say words rather than letters as they typed.

## Methods

**Subjects.** In experiment 1a, 50 subjects were recruited from AMT and compensated \$0.25 for participating in the approximately 5 minute task. 40 subjects completed the task



and were included in the analysis. In experiment 1b, 16 subjects were recruited from the undergraduate population at Brooklyn College, and received course credit for their participation. One subject was not included in the analysis because they repeatedly typed “asdf”, rather than completing the task.

**Apparatus & Stimuli.** For experiment 1a, typing tests were programmed for the online environment using HTML and JavaScript and conducted in subjects’ web-browsers. For experiment 1b, typing tests were conducted on an iMac (21" screen) controlled by LIVECODE software. Typing responses were registered on a standard QWERTY keyboard. Each paragraph in the typing task involved copy-typing short paragraphs (~115 words length), taken from Logan and Zbrodoff (1998).

**Design and Procedure.** During the task, participants were shown each paragraph in a text box. Paragraph text was black and presented in 14pt, Helvetica font. In general, participants were instructed to begin typing with the first letter in the paragraph. Correctly typed letters turned green, and typists could only proceed to the next by typing the current letter correctly. For each of the different typing conditions, subjects were presented with the following instructions.

For experiment 1a, the instructions read: This experiment tests how your inner voice influences your typing ability. When you type this paragraph, use your inner voice to say each word (or letter in the letter condition) that you type. At the beginning of the experiment, subjects were also asked to choose whether they use their inner voice during typing to speak either words or letters.

For experiment 1b, in the speak aloud condition, the instructions read: When you type this paragraph, speak aloud each word (or letter in the letter condition) that you type. In the inner voice condition, the instructions were identical to experiment 1a. At the beginning of the experiment, subjects were given a short questionnaire about how they use their inner

voice during typing. The questionnaire stated “People have the ability to use their inner voice to speak silently to themselves. We are interested in how you use your inner voice while you are typing. Please estimate the percentage of times that your inner voice is engaged in each of the following tasks while typing, and make sure your numbers add up to 100.”. The four tasks were: a) inner voice is silent, b) inner voice speaks words that you are typing, c) inner voice speaks letters that you are typing, d) inner voice speaks other words or letters that you are not typing.

## Results

**Experiment 1a.** For each subject, we applied the following pre-processing steps. We included IKSIs only for keystrokes involving a lower case letter, and only for correct keystrokes that were preceded by a correct keystroke. Outlier IKSIs were removed for each subject, on a cell-by-cell basis, using the Van Selst and Jolicoeur (1994) non-recursive moving criterion procedure, which eliminated approximately 0.05 of IKSIs from further analysis. An alpha criterion of .05 was set for all analyses.

Mean IKSIs for correctly typed letters were computed for each subject in each condition and submitted to a 2 (Linguistic Unit: Say Letters vs. Say Words) x 2 (Letter position: First letter vs. Middle Letter) repeated measures ANOVA. Mean IKSIs are displayed in Figure 1.

The main effect of the linguistic unit condition was significant,  $F(1, 39) = 24.89$ ,  $MSE = 3,082.97$ ,  $p < .001$ ,  $\hat{\eta}_G^2 = .039$ . Mean IKSIs were shorter when subjects were instructed to use their inner voice to say the words (170), rather than the letters (214) they were typing.

The main effect of letter position was significant,  $F(1, 39) = 22.28$ ,  $MSE = 6,784.80$ ,  $p < .001$ ,  $\hat{\eta}_G^2 = .074$ . Mean IKSIs were longer for letters in the first position (223), compared to letters in the other positions (161).

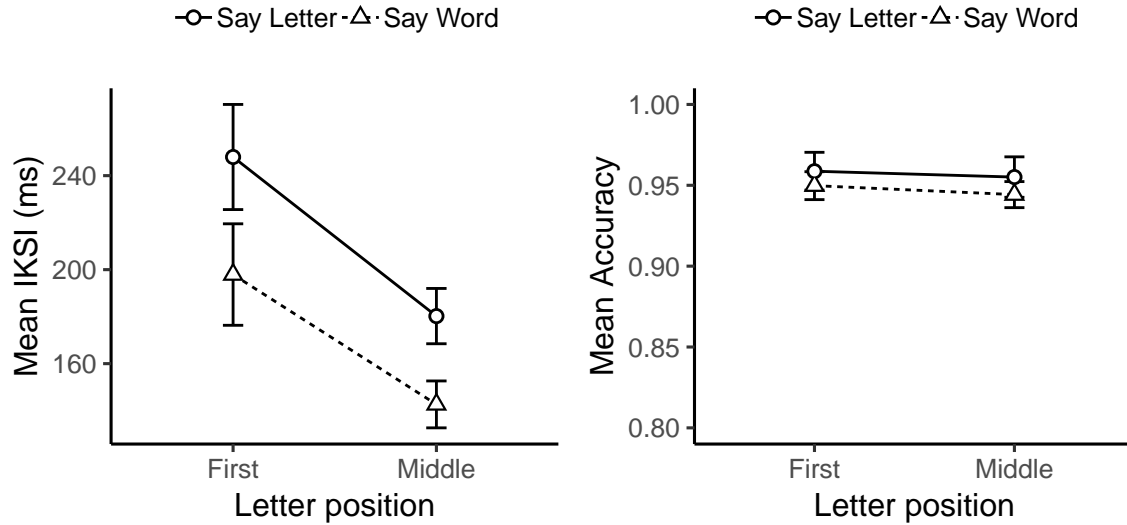


Figure 1. Mean interkeystroke intervals (IKSIs in ms) and mean accuracy as a function of linguistic unit condition (Say Word vs. Say Letter), and letter position in word (first letter vs. middle letter). Middle letters were any remaining letter in the word.

The interaction between linguistic unit and letter position did not meet the significance criterion,  $F(1, 39) = 3.08$ ,  $MSE = 501.77$ ,  $p = .087$ ,  $\eta_G^2 = .001$ .

A corresponding 2x2 repeated measures ANOVA was conducted on the mean accuracies for typing letters in each condition. Mean accuracies are shown in Figure 1. Accuracy was uniformly high in all conditions, and there were no main effects or interaction.

Last, prior to the experiments subjects were asked if their inner voice usually thinks in words or letters when they type. The proportions of subjects reporting thinking in words, letters, or who did not answer were 0.9, 0.08, 0.02, respectively.

**Experiment 1b.** Similar pre-processing rules were applied to the data. The exception was that accuracy was calculated at the word level rather than the letter level because the lab version of the experiment did not record accuracy at the letter level. Instead, IKSIs from words that were typed correctly were included in the IKSI analysis. The same outlier elimination procedure was applied, which excluded an average of 0.06 observations

from each condition. The accuracy analysis was conducted at the word level. Mean IKSIs and word level accuracy for each condition are displayed in Figure 2.

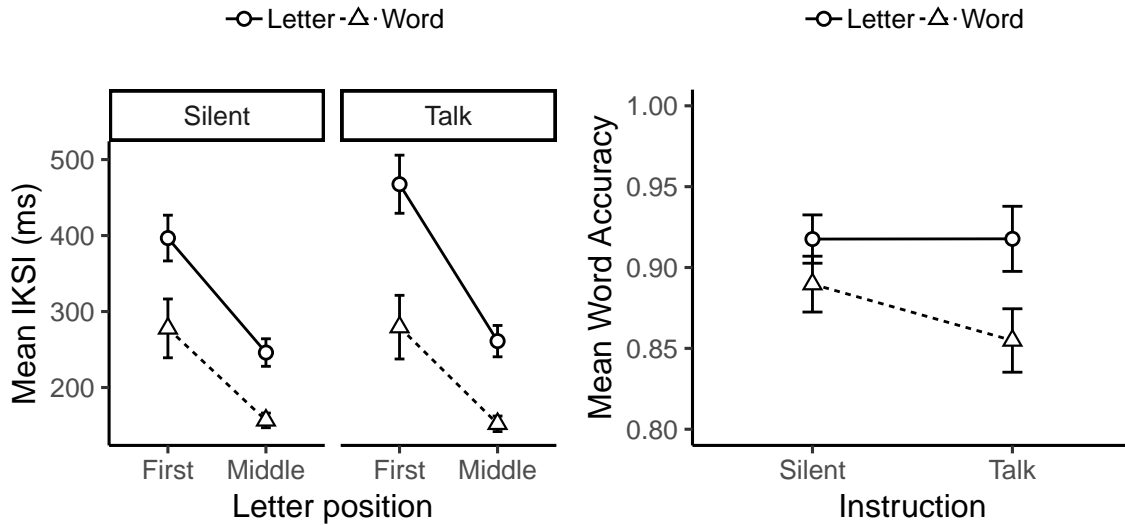


Figure 2. Mean interkeystroke intervals (IKSIs in ms) as a function of voice (silent vs. talking), linguistic unit (Say Letter vs. Say word), and letter position (first vs. middle); and, mean accuracy as a function of voice and linguistic unit.

Mean IKSIs for correctly typed letters were computed for each subject in each condition and submitted to a 2 (Voice: Silent vs. Talk) x 2 (Linguistic Unit: words vs. letters) x 2 (Letter position: first vs. middle) repeated measures ANOVA.

The main effect of the linguistic unit condition was significant,  $F(1, 14) = 24.45$ ,  $MSE = 19,560.37$ ,  $p < .001$ ,  $\eta_G^2 = .257$ . Mean IKSIs were shorter when subjects were instructed to use their inner voice to say the words (217 ms), rather than the letters (343 ms) they were typing. The main effect of letter position was significant,  $F(1, 14) = 24.02$ ,  $MSE = 28,607.35$ ,  $p < .001$ ,  $\eta_G^2 = .332$ . Mean IKSIs were longer for the first letter (355 ms), compared to other letters (161 ms). The main effect of voice was not significant, . Mean IKSIs were shorter for silent speaking (269 ms), compared to talking out loud (290 ms). These main effects were qualified by two interactions.

The voice by letter position interaction was significant,  $F(1, 14) = 4.66$ ,

$MSE = 1,544.69$ ,  $p = .049$ ,  $\hat{\eta}_G^2 = .005$ . The first letter slowing effect was larger when subjects spoke out loud rather than with their inner voice. Similarly, the linguistic unit by letter position interaction was significant,  $F(1, 14) = 7.82$ ,  $MSE = 2,846.27$ ,  $p = .014$ ,  $\hat{\eta}_G^2 = .016$ . The first letter slowing effect was larger when subjects said the letters rather than the words as they typed. The three-way interaction was not significant,  $F(1, 14) = 3.28$ ,  $MSE = 1,410.80$ ,  $p = .091$ ,  $\hat{\eta}_G^2 = .003$ .

The analysis of accuracy at the word level showed only a main effect of linguistic unit,  $F(1, 14) = 17.96$ ,  $MSE = 0.00$ ,  $p = .001$ ,  $\hat{\eta}_G^2 = .101$ . Subjects typed words more accurately when they said letters (0.92) compared to words (0.87). The interaction between voice and linguistic unit was not significant,  $F(1, 14) = 1.46$ ,  $MSE = 0.00$ ,  $p = .247$ ,  $\hat{\eta}_G^2 = .016$ .

The mean responses for the inner voice questionnaire were that subjects' inner voice was silent 3%, said words 69%, said letters 16%, or said unrelated words or letters 12% of the time during typing.

## Discussion

The main finding from both experiments was that mean typing speed was much faster when subjects covertly or overtly said the words they were typing rather than the letters they were typing. The behavioral data was also supported by the questionnaire data which showed that most subjects claim to use their inner voice to say the words they are typing rather than say letters, unrelated words or letters, or nothing at all. This finding is consistent with the idea that verbal codes at the word level are important mediators of keystroke production during typing.

## Experiment 2: Verbal Suppression

Verbal suppression techniques are commonly used in memory studies to prevent subjects from using their inner voice to verbally code and rehearse to-be-remembered information. The technique usually involves asking subjects to repeatedly recite a word or phrase out loud during the rehearsal period. As a result, verbal processes are occupied by the reciting the word, and presumed to be unavailable for rehearsing information in working memory. Experiment 2 asked whether similar verbal suppression techniques would also disrupt typing performance. We had subjects copy-type one of six paragraphs under one normal condition, and five verbal suppression conditions. Each verbal suppression condition had typists say words or letters out loud while they typed. These conditions included: a) repeating the word “the”, b) repeating the words Tuesday and Thursday, c) repeating the letters of the alphabet in order, d) repeating randomly chosen letters, and e) counting backwards from X.

### Methods

**Subjects.** 15 subjects were recruited from the undergraduate population at Brooklyn College, and received course credit for their participation.

**Apparatus & Stimuli.** The same apparatus and stimuli from Experiment 1b were used in Experiment 2.

**Design and Procedure.** The same general typing test procedure from Experiment 1b was used, with the exception that typists copied six paragraphs. The instructions for each of the conditions were as follows:

Normal Typing instructions:

Say the instructions:

Say Tuesday/Thursday instructions:

Say alphabet instructions:

Say random letters instructions:

Count instructions:

## Results

Mean IKSIs for correctly typed letters were computed for each subject in each condition and submitted to a one-way repeated measures ANOVA with Suppression condition (Normal, The, Tuesday/Thursday, Alphabet, Random, and Count) as the sole factor. Mean IKSIs in each condition are displayed in table 2.

The main effect of Suppression was significant,  $F() = , MSE = , p < .$  Post-hoc tests showed that all mean IKSIs in all suppression conditions were significantly slower than the mean IKSIs in the normal typing condition. Additionally, X was X etc.

## General Discussion

## References

- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The Psychology of Learning and Motivation*, 8, 47–89.
- Botvinick, M. M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, 111(2), 395–429.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, Mass.: MIT Press.
- Liu, X., Crump, M. J. C., & Logan, G. D. (2010). Do you know where your fingers have been? Explicit knowledge of the spatial layout of the keyboard in skilled typists. *Memory & Cognition*, 38(4), 474–484. doi:10.3758/MC.38.4.474
- Logan, G. D., & Crump, M. J. C. (2009). The left hand doesn't know what the right hand is doing: The disruptive effects of attention to the hands in skilled typewriting. *Psychological Science*, 20(10), 1296–1300.
- Logan, G. D., & Crump, M. J. C. (2011). Hierarchical control of cognitive processes: The case for skilled typewriting. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 54, pp. 1–27). Elsevier.
- MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(3), 671–685.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97.



- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1–32.
- Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science*, 6(1), 1–36.
- Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A*, 47(3), 631–650.