Investigating Predictive vs. Random Task-Switching Using the CVOE Task

Nicholas P. Maxwell1, Jacob Namias1, & Mark J. Huﬀ1

* The University of Southern Mississippi

Author Note

Correspondence concerning this article should be addressed to Mark J. Huff, 118 College Dr, Hattiesburg, MS, 39406. E-mail: [mark.huff@usm.edu](mailto:nicholas.maxwell@usm.edu). [OSF NOTE HERE]

Abstract

Abstract will go here. . .

*Keywords:* Task-Switching; CVOE; Attentional Control; Working Memory

Investigating Predictive vs. Random Task-Switching Using the CVOE Task

Selectively attending to relevant information within one’s environment is critical for successfully staying on task when engaging in goal-directed behavior. Attentional systems play a key role in this process, as individuals with strong attentional control are more likely to ignore highly salient but unrelated information within their environment that would otherwise produce distractions. Traditionally, researchers have investigated attentional control by presenting participants with situations in which task-related information is contrasted with information that is highly salient yet is unrelated to the task at hand. These studies consistently show that when participants are required to actively suppress unrelated information, both response times (RTs) and error rates are increased. A prominent example of this effect was originally reported by Stroop (1935). In his seminal study investigating the effects of color naming, Stroop had participants read lists of color words, which had been printed using ink that was either congruent with the word (i.e., “Blue” printed in blue ink) or incongruent (e.g., “Blue” printed in green ink). Participants were instructed to quickly name the color ink in which the word had been printed, rather than reading the word. Thus, the successfully completing the Stroop task required participants to suppress the highly salient but task-unrelated information contained in the word.

Overall, both RTs and error rates were increased for color words presented in an incongruent ink (e.g., “Blue” presented in green ink) relative to when a congruent ink was used (e.g., “Blue” presented in blue ink). Often, the decreases in performance that are observed for incongruent trials are further exaggerated whenever an incongruent trial immediately follows a congruent trial (i.e., congruency sequence effect, CSE; Aschenbrenner & Balota, 2019, Egner, 2007). Thus, participants are slower to respond and provide less accurate responses whenever they must suppress task irrelevant information to successfully complete a task, and these deficiencies are further compounded whenever participants switch between multiple trial types.

The Stroop color naming task has received significant attention in the literature and has often been used as a measure of attentional control (see MacLeod, 1992). This is because to complete the task, individuals must successfully activate and maintain the task goal of naming the ink color while suppressing highly salient but task irrelevant information (reading the word). As a result, researchers interested in the effects of both healthy and unhealthy aging on attentional control processes have commonly used variations of the Stroop task as a means for assessing declines in attentional process that occur a function of aging. For example, Spieler, Balota, & Faust (1996) showed that overall performance on the Stroop task decreased as a function of both healthy aging and Alzheimer’s Disease (AD) diagnosis. Specifically, compared to younger adults, healthy older adults showed slower RTs (but not an increase in error rates). For AD individuals, however, large costs to both RTs and error rates were observed, even when these individuals were matched to healthy older adults of the same age. More recently, Hutchison, Balota, & Ducheck (2010) showed that the Stroop color naming task could be used to discriminate healthy aging from AD, suggesting that this task is sensitive to the breakdowns in attentional control that inherent to AD. It is evident, therefore, that attentional control is critical for keeping internal goals active, as participants with impaired attentional control systems experience greater difficulty staying on task when required to surpass task-irrelevant information.

While researchers have commonly used the Stroop task as a measure of attentional control, there has been and increased focus on using task-switching paradigms as an additional technique to investigate questions related to attentional control processes. In a standard task-switching paradigm, participants alternate between completing a set of contrasting tasks (e.g., Jersild, 1927; Rogers & Monsell, 1995). Task-switching paradigms typically present participants with at least two types of experimental conditions. First, participants complete *pure blocks* which focus exclusively on one task (i.e., participants only complete addition problems). Next, participants complete *switch blocks* in which they must quickly alternate between two contrasting tasks (i.e., addition on trial one but subtraction on trial two). Like the Stroop task, switch blocks require participants to attend to a relevant task-set (i.e., the current task instructions) while suppressing irrelevant but salient information from the inactive task-set. To assess differences in attentional control, response times (RTs) and error rates are compared between the two block types. Overall, studies investigating task-switching have repeatedly shown that participants commit more errors and have slower RTs for switch trials compared to non-switch trials, and, like the Stroop task, declines in attentional control can exaggerate these costs (e.g., Huff, Balota, Minear, Aschenbreener, & Duchek, 2015).

Although researchers have developed a variety of tasks to investigate task-switching effects, the present study focuses specifically on task-switching paradigms in which a direct comparison can be made between local and global switch costs (e.g., Huff et al. 2015; Mayr, 2001; Minear & Shah, 2008; etc.). These paradigms initially present participants with a set of pure blocks (one corresponding to each task-set). These pure blocks are then immediately followed by one or more switch blocks in which switch and non-switch trials are interleaved (e.g., switch, non-switch, switch, non-switch, etc.). The *global switch cost* refers to the difference between switch trials and pure block trials and represents the cost of maintaining multiple task configurations in a switch block compared to a single task configuration within the pure block (Minear & Shah, 2008; Wylie & Allport, 2000). Alternatively, the *local switch cost* is found by computing the difference between switch and non-switch trials presented within the same switch block. Local costs represent task-set reconfiguration processes that occur due to participants changing tasks-sets within the same block of trials (Rogers & Monsell, 1995; see Huff, et al., 2015).

While declines in attentional control due to aging have been shown to influence task-switching performance (see Huff et al., 2015), the stimuli used have also been shown to influence both accuracy and RTs. For example, previous research has shown that switch costs are particularly magnified whenever the stimuli do not clearly signal to participants which of the two tasks is to be performed (Luwel, Schillemans, Ongehan, & Vershaffel; 2009). Termed *bivalent* stimuli, these items activate both task-sets used in a switch task (i.e., presenting participants with letter-number pairs and having them switch between classifying the letter or the number). Compared to *univalent* stimuli which only correspond to one task-set (i.e., presenting participants with only letters or numbers), responses are often slowed, as participants must consider both task sets (e.g., bivalency cost; Woodward; Meier, Tipper, & Graf; 2003). Because bivalent stimuli are more challenging for participants, researchers have often commonly incorporated them when developing task switching paradigms. One example is the Consonant-Vowel Odd-Even switch-task (CVOE; Minear & Shah, 2008), which is a classification task that presents participants with letter-number pairs (e.g., A 15). Depending on the task-set being cued, participants are instructed to either classify the letter in the pair as being a consonant/vowel or the number as being odd/even. Because this task presents participants with a set of pure blocks before the switch block, the CVOE task therefore allows for measurement of both local and global switch costs.

Within the past decade, researchers have made extensive use of the CVOE task to investigate questions related to attentional control. Often, studies using the CVOE task have been interested in the effects of healthy and unhealthy aging on attentional control. For example, Tse, Balota, Yap, Duchek, & McCabe (2011) compared performance between young, healthy older adults, and older adults with mild cognitive impairments (MCI) on three attentional control measures including the CVOE task. Though Tse et al. were primarily interested in distributional measures of RTs as a measure of attentional control processes (rather than traditional analyses of mean RTs and Error rates), we note that MCI individuals showed greater local switch costs for errors relative to younger adults. For RTs, MCI individuals showed decreased local costs compared to healthy older adults. Tse et al. attributed the increased cost to errors and the subsequent decrease in local costs to RTs as being due to MCI individuals having greater difficulty suppressing the inactive task set when switching.

More recently, Huff et al. (2015) compared CVOE task-switching between young adults, healthy middle aged and older adults, and MCI older adults. Overall, MCI older adults committed more errors and had slower RTs relative to both young adults and healthy older adults, with task performance particularly affected for switch trials compared non-switch trials in which the task-set does not change. Importantly, Huff et al. (2015) compared changes in global and local costs of both errors and RTs as a function of age and MCI status. First, global switch costs (switch trials compared to pure trials) for errors increased as a function of both age and MCI status. This pattern subsequently extended to global costs of RTs, suggesting that requirement to keep two task sets active placed additional burdens on working memory. For local costs, however, no group differences in errors emerged, but local costs of RTs decreased across groups, suggesting that MCI individuals were not as well tuned to the task set relative to younger adults and healthy older adults.

**Predictable vs Random Switching**

In addition to the type of stimuli (bivalent vs univalent), task-switching paradigms can be further classified based on whether switching occurs in a predictable or random sequence. A commonly used predictable switch paradigm is the *alternating runs* sequence (e.g., Rogers & Monsell, 1995; Huff et al., 2015). Participants completing an alternating runs switch task change task every *n* trials (e.g., AABBAABB, etc.). As a result of this sequence, participants are generally aware of when task-switches will occur. Alternatively, task switching may occur at unpredictable intervals. In this case, the upcoming task is unknown to participants until they are cued on the next trial. Random task switching can be further divided into task-cueing paradigms (e.g., Meiran, 1996) in which a cue is available at each trial or can employ intermittent instructions (i.e., random interruptions to change task-set; Gopher, Armony, & Greenshpan, 2000; see Monsell, Sumner, Waters, 2003 for a review of predictable vs random task-switching paradigms). [BENEFITS OF EACH PARADIGM]

[Overview of Monsell et al., 2003]

Regarding the CVOE task, [MINEAR AND SHAH]

[TRANSITION TO THE PRESENT STUDY] [TRANSITION SENTENCE TO RT DISTRIBUTIONS]

**Distributional Analyses of RTs**

Across task-switching studies, [EXPAND – SOMETHING ABOUT RTS BEING FUNDAMENTAL] Commonly, researchers have commonly relied upon analysis of mean or median RTs as a method to gain insight into attentional processes. However, because RT distributions are almost always positively skewed (i.e., most RTs generally occurring at the faster end of the scale), performing an analysis of only mean RTs may produce results that are misleading (see Balota & Yap, 2011 for a review). To account for this, researchers have increasingly moved away from using traditional measures of central tendency when assessing RTs, and instead, have elected to focus on RT distributions.

Previous research has shown that the characteristics of RT distributions can capture important aspects of human cognition, including word recognition (e.g., Andrews & Heathcote, 2001; EXAMPLE), semantic priming (e.g., Balota, Yap, Cortese, & Watson, 2008), and importantly, selective attention [CITE], and, importantly, attentional control [CITE]. [TRANSITION] For example, Tse et al. (2011) [SUMMARY] More recently, [HUFF ET AL HERE] [CONCLUSION STATEMENT HERE]

[BREAKDOWN THE ATTENTION STUDY FINDINGS HERE]

Given the benefits of using RT distributions to investigate attentional control processes, the present study incorporates these distributional analyses in addition to traditional mean analyses. In doing so, we include two types of RT distributional analyses: 1) Averaging RTs across participants and binning them via a Vincentile analysis and 2) fitting individual RTs to an ex-Gaussian distribution. First, the Vincentile analysis rank orders all RTs for each trial type at the participant level and then bins the ordered data into groups of equal size. For example, a Vincentile analyses using four bins would first each participant’s RTs from fastest to slowest. Next, for each participant, RTs within the first 25% of the data would be averaged, followed by the second 25%, third the 25%, and the final 25%. This process is then repeated for each participant, and Vincentiles are computed by taking the average of each bin across participants. [TALK ABOUT THE SHAPE OF THE DISTRIBUTION?] [NEED TO EXPAND SOMEHOW]

The ex-Gaussian analysis [EX-GAUSS HERE]

**The Present Study**

Given differences in switch costs that could potentially arise due to differences in stimuli presentation sequences, the goal of the present study was to expand upon previous research on CVOE task-switching by comparing error rates and RTs between participants who complete switch blocks in which task-switching occurs via a predictable alternating runs sequence (e.g., CV-CV-OE-OE-CV-CV) and participants who receive switch blocks in which task-switching occurs at random (e.g., CV-OE-OE-OE-CV-OE). Overall, we expected that mean error rates and RTs would be higher on switch blocks (regardless of presentation sequence) relative to pure blocks, given that pure blocks only require participants to engage in a single task-set. However, within switch blocks, we expected that participants would particularly struggle whenever switching occurred at non-predictive intervals due to the lack of a discernable pattern. We therefore anticipated that the random switch block would produce more errors and higher RTs relative to the alternating runs switch block, and further, that these difficulties would produce greater local costs compared to alternating runs task-switching, as [WHY?] Finally, regarding global costs [GLOBAL COSTS PREDICTION]

**Method**

**Participants**

A total of 100 undergraduate students were recruited from the University of Southern Mississippi’s undergraduate research pool and completed the study in exchange for partial course credit. Data from 9 participants were removed due to excessive error rates in either the pure or switch blocks (i.e., mean error rates within a block that were > 3 standard deviations above the mean), which indicated that participants did not correctly follow task instructions. Additionally, data for two participants were removed due to a coding error. A sensitivity analysis conducted with *G\*Power* [CITE] indicated that our final sample of 89 participants was sufficient to detect XX effects [STATS]. All participants were native English speakers who reported normal or corrected to normal vision. Participant demographics are reported in Table X.

**Materials**

To create the stimuli, we generated a series of letter-number stimulus pairs (e.g., A 15) using the following process. First, an even number of consonants and vowels were created. These letters were always selected from A, D, E, H, I, J, O, P, S, or U. Next, a series of numbers were randomly selected between 1 and 99, with the constraint that half of the numbers selected were always even. To create the pairs, half of the consonants were paired with an odd number, while the remaining half were paired with even numbers. This process was then repeated for vowels. This resulted in an equal number of each of the four possible stimulus pair types (Consonant-Odd, Consonant-Even, Vowel-Odd, Vowel-Even) within each block. Letters and numbers repeated within blocks, however, pairs were arranged within each block such that repeats did not occur on consecutive trials.

**Procedure**

The CVOE task presented participants with two sets of instructions, which either differed between blocks (pure blocks) or as a function of trial (switch blocks). For each trial, a letter-number pair was presented in the center of the computer screen, and participants were tasked with classifying whether the letter was a consonant/vowel (CV trials) or an odd/even number (OE trials). Depending on the type of trial, the words consonant/vowel or odd/even were presented at the top of the screen in the left and right corners to serve as a reminder. Participants were instructed to press the *q* key for consonants/odd numbers or the *p* key for vowels/even numbers. These keys were selected given that they are on opposites sides of a standard QWERTY keyboard. Stimuli were presented in 30-point Courier New font, and trials were presented with a 500 ms intertrial delay.

Trials were arranged into four blocks, with each block containing an equal distribution of *q* and *p* responses. Following the design of Huff et al. (2015), participants first completed two pure blocks (CV and OE) before completing two switch blocks (alternating runs and random presentation). Participants initially completed a set of 10 practice trials which corresponded to the first pure block’s task (CV or OE) and received verbal feedback on their performance. Following completion of the practice phase, participants immediately began the first pure block. Pure blocks each contained 96 trials and focused exclusively on one of the two tasks, with one block containing the CV task and the other the OE task. Following completion of the first pure block, participants completed a second set of practice trials (corresponding to the task in the second pure block) before completing the second pure block.

Immediately following completion of the two pure blocks, participants began the two switch blocks. In the switch blocks, the task change occurred at the trial level rather than the block level. For each trial, participants were prompted with the word “letter” or “number”, which corresponded to the CV or OE task, respectively. This prompt was located above the stimulus pair, and participants were informed that the prompt could potentially change following each key press. To practice the switching task and become familiar with the prompts, participants first completed a set of ten practice switch trials. Following this practice session, participants immediately began the first switch block. Trials within the switch blocks were arranged such that they were presented either with an alternating runs pattern (e.g., CV, CV, OE, OE, CV, CV, etc.; see Huff et al., 2015) or presented using a random presentation sequence (e.g., CV, OE, OE, OE, CV, OE, etc.). Each switch block consisted of 120 trials, which consisted of 59 switch trials (i.e., a CV trial followed by an OE trial) and 61 nonswitch trials (i.e., two consecutive OE trials). Like the pure blocks, each switch block corresponded to one of these two presentation modes (alternating runs or random). Thus, participants completed one pure CV block, one pure OE block, one alternating run switch block, and one random presentation switch block. Block presentation was randomized across participants; however, blocks were always ordered such that participants completed the two pure blocks before completing the two switch blocks (Huff et al., 2015; Minear & Shah, 2008).

Across blocks, participants were instructed to respond to each trial as quickly as possible without compromising accuracy. The task was presented to participants on a laptop running E-Prime 3.0 software [CITE], and all participants were tested individually in a laboratory setting. The total experiment lasted approximately 20 minutes.

**Results**

For all analyses, significance was set at the *p* < .05 level. Generalized-eta squared (*η*2G) and Cohen’s *d* effect size estimates were computed for all significant analyses of variance (ANOVAs) and *t*-tests, respectively. In addition to reporting effect size indices, we supplemented all standard null-hypothesis testing with a Bayesian estimation of the strength of evidence in favor of the null hypothesis, which compares a model that assumes a significant effect to one that assumes a null effect (Masson, 2011; Wagenmakers, 2007). This analysis returns a probability estimate termed *p*BIC (Bayesian Information Criterion) which represents the likelihood that the null hypothesis is retained. Therefore, all null effects include a *p*BIC estimate.

In the following analyses, we first examine mean error rates across trial types (pure, alternating switch, alternating nonswitch, random switch, and random nonswitch) and switch cost (local vs global). We then assess changes in mean RTs across trial types and switch costs. Following the design of Huff et al. (2015), all RT analyses only included correct trials. Additionally, we employed a trimming procedure to reduce the likelihood of RT analyses being disproportionately influenced by extreme scores. We defined RT outliers as any responses occurring three standard deviations above or below of each participant’s respective mean. As a result of this trimming procedure, xx% of pure block trials, xx% of nonswitch trials, and xx% of switch trials were eliminated. Finally, [DISTRIBUTIONAL STUFF HERE] For completeness, error rates and RT comparisons are reported in the Appendix (Tables AX and AX, respectively).

**Mean Error Rates**

Mean error rates as a function of trial type are displayed in Figure 1 (top panel). Overall, participants committed the most errors on alternating runs switch trials (6.12%), followed by random switch trials (5.17%), alternating runs nonswitch trials (3.49%), pure trials (3.25%), and random nonswitch trials (3.01%). A one-way repeated measures ANOVA confirmed that error rates differed as a function of trial type, *F*(4, 352) = 20.29, *MSE* = 8.16, *η*2G = .09. Post-hoc *t*-tests revealed that this effect was driven primarily by increased errors for switch trials relative to nonswitch and pure trials, *t*s ≥ 3.63, *d*s ≥ 0.43. For switch trials, mean error rates were marginally greater when trials were presented using alternating runs compared to random presentation, *t*(88) = 1.92, *SEM* = 0.50, *p* = .06, *d* = 0.21, *p*BIC = .60. However, no differences were detected between pure and nonswitch trials, regardless of switch block presentation, *t*s < 1, *p*s ≥ .48, *p*BICs ≥ .88.

Next, we compared differences in switch costs for errors as a function of presentation and cost type (Figure 1, bottom panel). A 2 (Switch Cost: Local vs Global) × 2 (Presentation: Alternating Runs vs Random) yielded a significant main effect of Switch Cost, *F*(1, 88) = 26.83, *MSE* = 19.03, *η*2G = .10, such that collapsed across presentation modes, local switch costs exceeded global costs (2.39 vs. -0.01). Additionally, this analyses revealed a marginal effect Presentation, *F*(1, 88) = 3.68, *MSE* = 5.43, *p* = .06, *p*BIC = .60, *η*2G = .01. The interaction between Switch Cost and Presentation, however, was not significant, *F*(1, 88) < 1, *MSE* = 17.35, *p* = .99, *p*BIC = .90.

**Mean RTs**

As displayed in Figure 2 (top panel), participants were quickest when responding to trials presented in pure blocks compared to switch and non-switch trials. A one-way repeated measures ANOVA confirmed the presence of trial type differences, *F*(4, 352) = 357.72, *MSE* = 19.03, *η*2G = .10. Post-hoc testing indicated that this effect was largely driven by differences in RTs between trials presented in the pure and switch blocks

**Vincentile Plots**

[VINCENTILES] [WILL NEED TO RUN ANOVAS]

**Ex-Gaussian Distribution of RTs**

[EX-GAUSS]

**General Discussion**

The goal of the present study was to [SUMMARY PARAGRAPH – MAIN ANALYSES] Overall… [EXPAND]

[SUMMARY PARAGRAPH – DISTRIBUTIONAL ANALYSES]

[SOMETHING HERE – I’LL FIGURE IT OUT LATER]

[AGING IMPLICATIONS]

[FUTURE DIRECTIONS]

**Summary and Conclusion**

[WORDS HERE]

**References**

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 514–544.

Aschenbrenner, A. J., & Balota, D. A. (2019). Additive effects of item-specific and congruency sequence effects in the vocal Stroop task. *Frontiers in Psychology, 10*, 860.

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science, 20*(3), 160-166.

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59*(4), 495-523.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience, 7*, 380-390.

Huff, M. J., Balota, D. A., Minear, M., Aschenbrenner, A. J., & Duchek, J. M. (2015). Dissociative global and local task-switching costs across younger adults, middle-aged adults, older adults, and very mild Alzheimer’s disease individuals. *Psychology and Aging, 30*(4), 727–739.

Hutchison, K. A., Balota, D. A., & Duchek, J. M. (2010). The utility of Stroop task switching as a marker for early-state Alzheimer’s disease. *Psychology and Aging, 25*(3), 545-559.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology, Whole No. 89*.

Luwel, K., Schillemans, V., Onghena, P., & Verschaffel, L. (2009). Does switching between strategies within the same task involve a cost? *British Journal of Psychology, 100*(4), 753-771.

MacLeod, C. M. (1992). The Stroop task: The “gold standard” of attentional measures. *Journal of Experimental Psychology: General, 121*(1), 12-14.

Mayr, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging, 16*, 96-109.

Minear, M. & Shah, P. (2008). Training and transfer effects in task switching. *Memory & Cognition, 36*, 1470-1483

Rogers, R. D., & Monsell, S. (1995). Costs of predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207-231.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22*(2), 461–479.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Tse, C.-S., Balota, D. A., Yap, M. J., Duchek, J. M., & McCabe, D. P. (2010). Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology, 24*(3), 300–315.

Woodward, T. S., Meier, B., Tipper, C., & Graf, P. (2003). Bivalency is costly: Bivalent stimuli elicit cautious responding. Experimental *Psychology*, 50(4), 233–238.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of “switch costs.” *Psychological Research, 63*, 212-233.

[TABLE 1]