October XXth, 2022

Jarrod Lewis-Peacock, PhD

Action Editor

*Attention, Perception, and Psychophysics*

Dear Dr. Lewis-Peacock,

We have submitted a revised version of our manuscript PP-BB-22-011 “Predictive Alternating Runs and Random Task-Switching Sequences Produce Dissociative Switch Costs in the Consonant-Vowel/Odd-Even Task.” We are encouraged that our reviewers found our study to be “interesting and novel” and, particularly, that Reviewer 2 felt that our study provides “a valuable contribution to the task-switching literature.” Below, we list our responses to each reviewer’s comments. To facilitate review, we cite page numbers when referencing specific changes and have made all primary modifications to the manuscript using blue-colored font. We hope that our revised manuscript is now suitable for publication in *Attention, Perception, and Psychophysics.*

Sincerely,

Mark J. Huff, PhD.

School of Psychology

The University of Southern Mississippi

Email: mark.huff@usm.edu

Phone: (601) 266-5411

Fax: (601) 266-5580

Cc:

Nicholas P. Maxwell, PhD

The University of Southern Mississippi

Jacob Namias, M.A.

The University of Southern Mississippi

**Action Editor (Jarrod Lewis-Peacock)**

**Comment 1:** One major concern from R1, that I agree with, is the appropriateness of this article for a special issue on working memory. At this time, it is unclear to me how this study advances our understanding of working memory. The sole mention of working memory here at present is for an interpretation of the increase in global switch costs for the predictive alternating runs, where it is suggested that working memory is taxed by "maintaining two task sets" and by the need to "monitor their progress across trials" to anticipate when a switch will be occurring. First, what is the evidence that both task sets were being actively maintained in working memory?

***Response:*** By nature, task-switching paradigms like CVOE require that participants repeatedly alternate between completing two different task-sets. For participants to successfully complete both tasks while still maintaining speed and accuracy, they must be able to keep both task-sets active in working memory to adjust when the task set changes. Research on individuals with impaired working memory systems suggests that working memory plays a critical role in task-switching performance. For example, older adults typically produce more errors and are slower at responding when completing task-switching compared to younger adults, whose attentional control and working memory systems are thought to be more intact (see Balota et al., 2010). These differences become even more exaggerated for older adults diagnosed with working memory impairments, such as Alzheimer’s Disease (e.g., Huff et al., 2015). The focus of our manuscript is how well working memory processes are affected by different tasks switching contexts. We have significantly revised the Abstract and the Introduction (e.g., pgs. 4-6) to clarify the contributions of working memory processes.

**Comment 2:** Second, what would this task-switching performance tell us about working memory?

***Response:*** As mentioned in our response above, task-switching has often been used to assess breakdowns in attentional and working memory processes due to both healthy and unhealthy aging (e.g., Balota et al., 2010; Huff et al., 2015; Tse et al., 2010). Regarding the present study, our manipulation of switch sequencing (i.e., random vs. predictive switching) gives insight into conditions that challenge working memory processes in healthy adults. Our finding that random switching increases local switch costs versus predictive switching indicates that participants struggle to maintain both task sets in working memory and are slower and more likely to produce an error when they must reconfigure to a changing task set that is unexpected. Relatedly, increased global costs for predictive switching (vs. random) likely reflect an additional burden on working memory due to participants monitoring their progress through a trial sequence in working memory while simultaneously keeping each task-set active. Thus, task-switching paradigms are excellent for gauging working memory processes both in terms of how participants switch between two task sets (i.e., local costs) and how participants maintain multiple tasks sets in working memory over repeated task sets (i.e., global costs).

**Comment 3:** In a potential revision, the authors would need to address all the concerns raised by the reviewers and would need to pay extra attention to the framing of this work for a special issue on working memory.

***Response:*** We have revised our manuscript to address each reviewer’s primary concerns while taking special care to clarify the relationship between task-switching processes and working memory (see pg. 4, Introduction and pg. 21, General Discussion for examples). In doing so, we now provide a more streamlined review of the literature in the Introduction (Reviewer 1, comment 1; Reviewer 2, comment 1) while expanding the contents this literature review to encompass a broader selection of task-switching studies as suggested by both reviewers. Additionally, we have clarified our methods by providing greater detail regarding trial timing (Reviewer 1, comment 5; Reviewer 2, comments 4 and 5, e.g., pg. 14). Finally, we have expanded our discussion of the underlying mechanisms driving task-switching processes while also clarifying the relationship between these processes and working memory (Reviewer 2, comment 2, e.g., pg. 23).

**Reviewer 1**

**Comment 1.** It appeared to me that the introduction and theoretical rationale of the study substantially lack focus. For instance, the authors describe Stroop effects across several paragraphs even though they are entirely irrelevant to the study at hand and the research question they want to address in the realm of task switching. Similarly, the authors mainly refer to task switching studies that compared specific participant subgroups rather than more general task switching studies. They comment on individual and group differences at several occassions although these differences also appear to be irrelevant to the study at hand. I might have missed something here, but in my opinion, the authors need to streamline the paper to get the storyline straight and more clearly and explicitly articulate their research question's theoretical relevance. At present, it appeared to me as if they just wanted to confirm specific data patterns that, based on my reading of the manuscript, do not really add anything to prior knowledge. Is that really the case and the study is just about confirming a data pattern nobody would have doubted or is there more to it?

Depending on what the authors actually want to argue for, they might also have to substantially adapt their General Discussion accordingly.

***Response:*** Thank you for your feedback. We agree that streamlining the focus of our introduction would provide greater clarity regarding the rationale behind our study. However, we disagree that a discussion of Stroop effects is irrelevant to the present research. Indeed, the Stroop task requires that participants suppress the automatic reading task by keeping the deliberate color-naming task-set active in working memory. Like CVOE switch task, the Stroop task requires that participants alternate between conflicting task-sets (i.e., reading word *colors* vs. reading word *names*) and both task sets must be maintained in working memory for successful completion. Additionally, participants must maintain the goal of reporting the color in working memory over the automatic word-reading response. Based on your suggestions, we have streamlined our discussion of Stroop effects to clarify the link between this task and other task-switching paradigms (pg. 4) while also relating Stroop to working memory processes. Our revised Introduction now places a greater emphasis on switch costs and their proposed mechanisms, including how these processes relate to working memory (e.g. pg. 6). In doing so, we now incorporate several of the references you suggested (see our response to Comment 3). We appreciate you providing us with this additional literature.  
  
**Comment 2:** Based on the submission information, it appears that the manuscript was submitted for a special issue in working memory. In this case, the authors do not seem to relate to the relevant literature on working memory and corresponding models of working memory at all and they either do not frame their research question appropriately and relate it to working memory or actually there is scarcely a relation. At least that was my understanding based on my reading of the current version of the manuscript.

***Response:*** Yes, you are correct that this manuscript was submitted for consideration in the special issue on working memory. We reasoned that this consideration was warranted, given that working memory processes are essential for successful completion of a switch task. In our revision, we have taken special care to highlight this link between task-switching processes and working memory (e.g., pg. 4; please see our response to the Action Editor, comment 1).

**Comment 3:** The authors present a very particular and limited set of task switching studies. They might want to, for instance, refer to more recent review articles on task switching (or some of the papers cited in them) and relevant underlying mechanisms in task switching related to their research question (once it is clarified):

***Response*:** Thank you for providing us with these additional references. We have updated our review of the task-switching literature in the Introduction to include this additional literature (e.g., pg. 6).  
  
**Comment 4:** On page 9 there seems to be a mistake when describing the results of Minear and Shah (2008). They report a pattern that is the exact opposite of their hypothesized results. That does not seem to be correct.

***Response***: We apologize for this discrepancy. We initially reported Minear and Shah’s pattern in our Introduction, given they previously included both a predictive and random CVOE switch block within the same study. This pattern is correctly reported on pg. 9 of our initial submission, and our current findings, while our data patterns were consistent with our hypotheses, they were not consistent with the sequence patterns reported by Minear and Shah.

We note, however, two discrepancies between our study and Minear and Shah’s. First, Minear and Shah were particularly interested in transfer effects on later switch task performance, rather than a direct comparison of predictive/random switching on task accuracy and switch costs. Second, and more importantly, it is unclear whether Minear and Shah counter-balanced across switch block type as was done in our experiments. Thus, all participants in their study may have first completed the predictive switch task before completing the random switch task, potentially leading to practice effects that influenced performance on the random switch block.

However, given the discrepancy between Minear and Shah’s reported data pattern and the present study, we have updated our predictions on pg. 12 to account for the possibility of replicating Minear and Shah’s pattern, while still retaining our initial predictions which are consistent with performance deficits anticipated under increased working memory loads. Additionally, we now address this discrepancy in the General Discussion (pg. 24).  
  
**Comment 5:** The methods should be described in more detail. For instance, the time course of a trial is not entirely clear. Was there a response deadline at all etc.? A figure of the trial time course would also be helpful in this regard.

Moreover, it was unclear to me, based on the authors description, whether participants were informed about the predictable versus unpredictable task sequence pattern. This is a rather important piece of information as the first trials of a task switching block should potentially not be included if participants were not informed about the predictable/unpredictable task sequence pattern beforehand.

Additionally, participant exclusions and outlier trial exclusions were not described clearly. What were the corresponding reference means here? For instance, were outliers determined based on the overall mean per participant or the individual cell means per participant (in my opinion the more appropriate option)?

***Response***: All trials were self-paced, and no response deadline was imposed. For each trial, participants were simply instructed to respond as quickly and accurately as possible. Based on your suggestion, we now include a figure detailing the time course within both pure and switch blocks (Figure 1, pg. 36). Concerning switch blocks, participants were not initially informed of the task-sequence pattern. However, as described in our initial submission, block order was counter-balanced across participants, such that some participants completed random switching before predictive switching and vice-versa. Thus, it is likely that participants’ a priori knowledge of the presentation sequence did not affect their subsequent task-performance.

Regarding outlier exclusions, we employed a trimming procedure based on Huff et al. 2015 in which, for each participant, mean RTs were computed across all trials. RT outliers were defined as any RT three or more standard deviations from the mean. Additionally, we removed RTs < 200 ms (which likely reflected anticipatory responses) and RTs > 10000 ms (which likely reflected participants being off task). Given this trimming procedure, it is likely that any differences in performance due to participants not anticipating random vs. predictive switching would be negligible.

We have updated the methods section to clarify trial pacing (pg. 15) and provide greater detail in the results section regarding the RT trimming procedure (pg. 17).

**Comment 6:** The authors present a rather generic sample size justification. They should at least provide information on the corresponding alpha level, power, and on which effect(s) their estimation refers to (Omnibus test of the univariate ANOVA vs. specific comparisons etc.).

***Response:*** We have updated the sample-size justification on pg. 13 accordingly.

**Comment 7:** The authors need to include information on their decision criterion regarding Bayes factors. According to the Bayesian criteria and analyses I am familiar with, they Bayes factors constitute inconclusive evidence if I am not mistaken. Here, a clarification of the meaning of certain value ranges is essential and should be added.

***Response:*** We have updated our results section (pg. 16) to provide additional information on what *p*bic represents (e.g., a probability estimate which asymptotes at 1.0 and corresponds to the likelihood of retaining the null which is highly sensitive to sample size, etc.). Interestingly, the recent movement towards Bayesian analyses was, at least in part, to avoid cutoffs that were present in standard NHST approaches. Ironically, some researchers report Bayes factors, which compute estimates at different “strength levels” which cutoffs are then applied to interpret these factors as either inconclusive, anecdotal, moderate, strong, and very strong for evidence of the alternate and null hypotheses (e.g., Kass & Rafferty, 1995). A strength of *p*bic is that does not use arbitrary cutoff scores and merely provides a probability estimate regarding the reliability of the null effect reported.

**Comment 8:** The authors conducted multiple analyses on the same RT data. It was not mentioned whether they correspondingly adjusted the alpha level. This information needs to be added. In addition, it was not clear to me why they authors did not only conduct the analysis including vincentiles rather than both the simpler ANOVA and this analysis.   
  
***Response:*** As mentioned on pg. 16 of our initial submission, alpha was set at .05 for all analyses. This is consistent with previous studies which have employed distributional analyses to assess RT data (e.g., Huff et al., 2015; Tse et al., 2010, etc.).

Regarding our inclusion of both distributional analyses and traditional ANOVAs, we elected to report both types for completeness, as we believe both types of analyses yield important information regarding mean RT differences and RT switch costs. This is consistent with other studies utilizing both traditional and distributional analyses when analyzing RT data (e.g., Balota & Yap, 2011; Huff et al., 2015; Speiler et al., 2000; Tse et al., 2010, etc.).

We appreciate you taking the time to review our manuscript.

**Reviewer 2**

**Comment 1:** Overall, I think the current review of literature on attention control in the Stroop task in the introduction could be trimmed down, and greater emphasis could be placed on the relevant task-switching literature. For example, the introduction section contains several paragraphs describing attentional control in Stroop task and age-related differences demonstrated when using this task. Since the Stroop Task was not used in the study, such a detailed description of this paradigm distracts the reader from the actual experimental task and research questions. Probably, the authors reviewed this literature to introduce attentional control and related control processes such as activation and maintaining the task goal while suppressing task-irrelevant responses. In this case, it would be enough to explain this point in several sentences, or to review articles that used task-switching designs instead. Further, the detailed reviewing of literature on age effects (e.g. p.7) creates an expectation that the authors would later report results for different age groups in their study as well, which is not the case. Because only young participants were tested in the study, the discussion of age effects on attentional control in the introduction is not particularly relevant for the research questions and a bit misleading.

***Response:*** Thank you for your feedback. Yes, you are correct that our Introduction includes a discussion of Stroop effects as a means of introducing attentional control processes. We have streamlined this section to be more in-line with the present research (please see our response to Reviewer 1, comment 1).

Similarly, in our initial submission, we included a discussion of aging effects in the Introduction, given that prior research utilizing the CVOE has often utilized this task to investigate the effects of aging on both task-switching processes and working memory (e.g., Huff et al., 2015; Tse et al., 2010). While we retain this section on pgs. 7-8, we have reworked it to provide a more focused discussion of both task-switching processes and their link to working memory (see our response to the Action Editor, comment 1).

**Comment 2:** Considering, that the primary goal of the study was to examine the effects of predicted vs. random task switching procedures on different types of switch costs, the article would benefit from describing the relevant processes in more detail in the introduction or at the latest in the discussion section using relevant task-switching literature. The key explanation of the results is a task-set reconfiguration process for elevated local switch costs in random task-switching, and keeping track of the task sequence for elevated global switch costs in alternating runs switching. However, the concept of task-set and its reconfiguration when switching from task to another are only shortly mentioned in the introduction. In the discussion section, some theoretical explanations are included, but in several places the references are lacking (e.g. p. 20, line 47; p. 21 lines 20-33). I would recommend the authors to better adjust the content of the introduction to their research questions and to try to better integrate their results into task-switching literature in the discussion (e.g. Andreadis & Quinlan, 2010; Shahar & Meiran).

***Response:*** Agreed. We have updated the Introduction to provide a more in-depth discussion of the cognitive processes involved with task-switching and have expanded our description of task-set reconfiguration processes (pg. 6). We have similarly updated the General Discussion to provide a more appropriate framing of our research questions and now include the suggested references (e.g., pg. 23). We appreciate you bringing this additional literature to our attention.

**Comment 3.** The approach to investigate the processes during switching task sets by analyzing the RT distributions is very laudable and represents an important and interesting line of evidence for distinguishing between different types of control processes. The approach to analyse vincentized RT distributions was proposed earlier by Ritske DeJong (2000) AN INTENTION-ACTIVATION ACCOUNT OF RESIDUAL SWITCH COSTS. Attention & Performance XVIII (Ed. Monsell and Driver). The authors should cite that article and relate their approach to that earlier work.   
  
In addition, a more careful consideration of earlier work is necessary throughout the manuscript. For example, the usage of the terms global and local switch costs seems at bit old fashioned because more recent work uses the terms mixing costs and switching costs. If the authors do not want to change their terms, then this is fine with me. However, in that case at some place a reminder would be helpful to inform that global switch costs are the same as what is usually termed mixing cost and local costs are termed usually as local costs.

***Response:*** We appreciate you bringing this additional literature to our attention. We now reference the De Jong chapter in several places within our manuscript (e.g., pg. 23).

Regarding the terminology used to describe switch costs, we initially selected the terms “local” and “global” to stay consistent with previous research that has employed these terms (e.g., Huff et al., 2015; Nashiro et al., 2018; Tse et al., 2010; Velichkovsky et al., 2020). Because most of these papers were published within the last decade, we believe these terms should not be characterized as old-fashioned. Additionally, we believe these terms provide greater clarity when discussing switch costs. Both “mixing costs” and “switching costs” reflect different types of task-switching costs, yet both types are technically switch costs. We therefore argue that local and global switch costs are more accurate terms.   
  
**Comment 4:** In the description of procedure, the information on how task-sets were cued in switch blocks is a bit confusing. On page 13 (lines 26-28), the authors write that “…the words consonant/vowel or odd/even were presented at the top of the screen in the left and right corners to cue the task set”. On the next page (lines 13-15), however, it is indicated, that “…participants were prompted with the word “letter” or “number….This prompt was located above the stimulus pair”. Do “cues” and “prompts” mean the same, namely to signal the participant which task has to be performed on a given trial, or are these different kinds of stimuli both presented in each trial of switch blocks? The authors should give a more precise description of the task design´.

***Response:*** By “cue the task-set,” we simply mean that the words consonant/odd and vowel/even were always presented in the upper left and upper right corners, respectively, to remind participants of which key was associated with each response (i.e., consonant/odd responses were made by pressing the “p” key, while vowel/even responses were made via the “q” key. For switch blocks, an additional cue was placed over the stimulus (e.g., “Letter” or “Number”), informing participants of which task was to be completed for a given trial.

We have updated our description of the study procedure (pg. 14) to clarify this point.  
  
**Comment 5:** It is also desirable if the exact timing of the trial would be specified. Even though, in this study, the CSI was not manipulated, the indication of how long “the prompt” (or “the cue”) and the test stimulus were presented, as well as the temporal sequence of the cue and the test stimulus, would contribute to the comparability of this study with other task-switching studies.   
  
***Response:*** All trials were self-paced, and participants were instructed to respond as quickly and accurately as possible. For switch blocks, prompts were displayed simultaneously with the stimuli. We now include a figure detailing trial time course for both pure and switch blocks (Figure 1, pg. 36; please see our response to Reviewer 1, comment 5.).

**Comment 6:** With regard to the results, the authors should be more cautious, when interpreting the dissociation between local and global switch costs as a function of the presentation sequence. The difference in switch costs is for the greater part driven by higher RTs in non-switch alternating runs trials. This supports the idea of tracking the trial sequence, which puts an additional load on working memory and leads to higher global costs, particularly in alternating runs paradigm. However, longer RTs in these trials also lead to smaller local switch costs in alternating runs, which restricts the interpretation that larger switch costs in random switching are due to more difficulty to reconfigure task-sets on switch trials. As a support for this interpretation, one would expect significant differences in switch trials between the two presentation sequences, which was not the case. Since the RT difference is only descriptive, I think, that this limitation is worth mentioning and discussing.

***Response:*** This is a fair point and one we now discuss as a potential limitation in the General Discussion (pg. 22).  
  
**Comment 7:** In addition, with regard to the Vincentized RT distribution for local switching costs as illustrated in Figure 2 and Figure 1. Did the authors use cumulative distribution functions for illustrating the vincentized RT distributions or not. If yes, then it is not clear to me how the RTs in the 6th bin (and perhaps in the 5th bin) in Figure 2 can be lower than the RTs in the bins 1-4. In case of cumulative distributions functions the RTs should be definitively longer in larger bins than in smaller bins. Otherwise the authors should include a reminder that the current analyses are not based on the common way of analyzing vincentized RT distributions because those usually are based on using cumulative data.

***Response:*** The vincentile analyses in Figure 2 were not cumulative and, rather, were conducted on switch costs. The finding that switch costs decrease on later bins simply indicates a smaller difference between trial types. Trial level vincentiles are reported in Figure 1.  
  
**Minor comments:**

**Comment 8:** p. 6 (lines 42-50): When explaining the CVOE task, the authors argue that it can be presented in pure and in switch block which allows the computation of the local and global switch costs. This passage sounds like a special advantage and therefore an argument for using this task. As far as I know, this possibility is not that much limited to the CVOE task: there are very many other tasks, that can be used in both pure and in switch blocks and allow the computation of different types of costs.

***Response:*** Yes, these switch cost types can be computed with other tasks. However, an advantage of the CVOE task is that the two task types used (consonant/vowel classification and odd/even classification) are relatively balanced in terms of participant knowledge (i.e., participants are typically not markedly better at one task set than another). If tasks were unbalanced, this would add additional variability when computing switch costs. This is an important point to note, but it is also important to highlight the benefits of using CVOE (pg. 7). We have updated this section accordingly and now reference other switch-tasks which similarly allow for pure block-switch block comparisons (pg. 6).

**Comment 9:** p.20 (lines 45-54): The paragraph is difficult to understand, even after repeated reading. The authors explain the finding that random switching increased local costs by referring to the study of Huff et al. (2015). Did they accidentally mention only the alternating runs condition? They further explain that younger adults produced larger local switch costs than older adults and why, and close the paragraph with the conclusion that local costs in random switching become exaggerated relative to predictive switching etc. The logic of this paragraph needs to be checked.

***Response:*** Huff et al.’s (2015) study did not include a random-switch block and thus the comparisons we make between the two switch block types were not available. We have clarified this in the paragraph. Huff et al. argued that the increased local costs on younger adults likely reflected a task-set inertia, which referred to a carry-over effect of the previous task set when the instructions shifted. Regarding local costs for random switching in the present study, we reasoned that if local costs reflected task-set inertia, these costs should be further elevated when task-set inertia is combined with random switching, which is inherently more difficult than predictive switching. Thus, random local costs are inflated relative to predictive local costs due to a combination of task-inertia and greater burdens placed on task-set reconfiguration processes. We have reworked this section on pg. 22 to clarify this point.

**Comment 10:** p.22 (lines 26-29): The authors discuss why a decrease in local switch costs across bins found in their study is in contrast to the results of Huff et al. (2015) who found an increase in these costs. They discuss the inclusion of additional trials and potential fatigue effects as one reason for this discrepancy in the results. This interpretation is not really plausible to me, since the occurrence of fatigue effects is usually interpreted as causing more lapses of attention, which should lead to longer RTs in the longest RT task trials. Here the authors seem to argue the opposite. At this point I would like additionally remind the discussion of the DeJong account (2000) for interpreting residual switch costs.

As an alternative, it might be possible that the decreasing size of local switch costs in the current study is caused by learning effects. Since the authors have presented twice as much trials as the study of Huff et al. it might be the cause that the relative and absolute number of switch trials was far higher in their study than in the study of Huff et al. This might have changed the state of preparation for a switching trial, which might be far higher in the current study than in the Huff study. I am not sure what may be the reason for the discrepancy in findings in these studies. However, the author should come up with a more plausible interpretation than the current one.

***Response:*** This is a fair point. We initially focused on fatigue effects, given that an obvious difference between our study and the one conducted by Huff et al. was our inclusion of both additional trials in each block and the inclusion of an additional, random switch block. We have revised this section on pgs. 23-24.

Regarding your second point on learning effects, we now mention this as a possible alternative explanation in the General Discussion (pg. 24)

All other minor grammatical and spelling errors have been corrected. We thank you for taking the time to review our manuscript.