**Action Editor**

I regret to inform you that the reviewers have raised serious concerns. R2 and R3  think that the paper is unsuitable for publication in the Quarterly Journal of Experimental Psychology. R1 has expressed major concerns too.

These substantial issues cannot be resolved in a single straightforward revision, so I am rejecting this paper.  It may be more suited to a more specialized journal, but QJEP will not be able to consider a revision of the manuscript.  
  
Sincerely,  
Barbara Treccani  
barbara.treccani@unitn.it  
  
 **Reviewer: 1**  
  
I would like to prepend that I am not well-versed in the task-switching literature.  
  
**#General Evaluation:**  
The authors present an investigation of the effect of the predictability of the task sequence on global and local task switch costs. They report that predictability increased the former (more typically referred to as ‘mixing costs’, to my knowledge) but reduced the latter, (at least descriptively) across all RT bins. Apart from details discussed below, the experimental procedure and statistical analysis appear sound. Some aspects of the data have not been explored, however, that should, in my view, be explored. Information gained from such an exploration may suggest a different interpretation than the one put forth by the authors, which does not quite convince me.

A general concern that I have is that the literature on the effects of the predictability of the task sequence on task switching may not have been discussed to the extent that it should be (e.g., Koch, 2007; Nessler et al., 2012). Even if an article does not report p-values on global or mixing costs because the experimental designs did not involve pure blocks, the descriptive effect of predictability on such costs can still be computed, or approximated from figures, since (unpredictable-non-switch-performance – pure-block-performance) – (predictable-non-switch-performance – pure-block-performance) = (unpredictable-non-switch-performance – predictable-non-switch-performance). In the case of the cited article by Monsell et al. (2003), for example, the figures show markedly better performance in the predictable case, indicating smaller mixing costs in that case. As I stated at the outset, however, I am not well-versed in this literature and there may be good reasons why the authors decided not to engage in a discussion of this kind.  
  
**#Ad Introduction:**(1)     The introduction could be shortened considerably, in my opinion. Especially the initial section about the Stroop task does not appear essential to me. If the authors insist that it should be included, the congruency sequence effect and the proportion congruency effect should be described such that it becomes apparent to the reader that they constitute interactions between congruency in the current trial and congruency in the previous trial and the proportion of congruent trials, respectively. The current descriptions could be misunderstood as referring to main effects.

(2)     Personally, I also do not think it is necessary to justify a distributional analysis in the introduction, especially not if it is performed in addition to more traditional analyses.  
  
**#Ad Method:**  
(1)     The estimated power value should be explicitly stated, not merely described as sufficient (Page 13).

(2)     The trial regime should be described more precisely. For example, did the stimulus remain on screen until the response? Was the letter always presented on the left and the number always on the right? (This might be slightly problematic, but might also afford an interesting opportunity [see final point])  
  
**#Ad Results/Analysis:**  
(1)     Since the authors used bivalent stimuli, I am surprised that they apparently did not investigate the (task) congruency effect.

A (task) congruent trial would be one in which the letter task requires the same response as the number task. Correspondingly, a (task) incongruent trial would be one in which the letter task requires a different response than the number task. This congruency effect can be regarded as a measure of the degree to which the cognitive system is able to prioritize the currently relevant task set over the currently irrelevant one (and suppress the latter).

There may be an interaction between this type of task congruency, switch/non-switch and predictability such that the difference between the congruency effects in switch vs non-switch trials (the former effect minus the latter) is less pronounced or potentially even reversed in the predictable scenario. After all, in the latter scenario, every non-switch-trial is a pre-switch-trial in which the cognitive system may already be in the process of preparing the upcoming switch and therefore reactivating the currently irrelevant task set.

This preparation of the switch could be the reason for the larger global costs, in the predictable scenario. I cannot convince myself that the monitoring of the progress through the trial sequence should have such an effect: As I understood the procedure, such a monitoring process is, strictly speaking, not even necessary because the task is cued in every trial any way, just as in the unpredictable scenario. Therefore, in my view, predictability should confer a benefit, rather than a cost.  
  
The degree to which one task set is prioritized over the other might also manifest in the Simon effect, especially if the letter/number is always presented on the same side of the screen: In this case, a congruent trial would be one in which the required response is on the same side as the stimulus (letter/number) to which it is made.  
  
In a potential future iteration, I would definitely like to see an analysis of at least the task congruency effect (as described above).  
  
Signed,  
Michael Sprengel  
(17.2.23)  
  
Reviewer: 2  
  
**Summary**  
In a single experiment, the authors compare performance of pure task blocks with a predictive and a random task switching condition. For RT data higher local switch costs (i.e., the difference between switch and nonswitch trials in mixed blocks) were found for random task switching, while higher global switch costs (i.e., the difference between nonswitch trials in mixed blocks with pure blocks) were observed for predictable switching. This result was interpreted in terms of more challenging task set reconfiguration in random switching and declines in task set maintenance for predictable switching, leading to the respective increase of switch costs.  
 **Evaluation**  
Overall, the manuscript is well written and easy to follow. However, from my point of view, the study adds only little new insights to the body of literature comparing random and predictable task switching. As other studies focusing on this topic have been conducted about 20 years ago, the present study looks a bit outdated. The use of Vincentile bins and of the Ex-Gaussian tau parameter in addition to RT and ER analyses is interesting but adds little additional information. Furthermore, I have some other major comments that I will elaborate below. As a consequence, I suggest to reject the manuscript.  
  
**Major Comments**  
The introduction focuses on performance differences between young versus old, or healthy versus impaired, subjects. As a consequence, the reader expects to find similar groups in the present study, However, only undergraduate students participated in this experiment – which is of course no general shortcoming, but this way, the introduction does not fit the rest of the study.  
  
Previous research (Milán et al, 2005, Acta; Monsell et al., 2003, Memory & Cognition) has found that RTs for task repetitions decrease after the first consecutive task repetition when switching was random. In the present study, the predictable switching condition included no more than one successive task repetition (as the sequence was A A B B A A etc.). However, although this is not explicitly mentioned, I guess that the random switching condition did include two or more successive task repetitions. This may have influenced the data pattern. If one compares the respective mean RTs (1328ms for predictable and 1260ms for random task repetitions), this really seems to be the case, so the data pattern is confounded.  
  
On p. 21, the authors argue that “unpredictable switch trials are particularly difficult” when explaining higher local switch costs in this condition. However, this reasoning is at odds with ER data, which were marginally lower for unpredictable switching. From my point if view, higher RT switch costs in the unpredictable condition are (at least partially) due to lower RTs for repetitions instead of higher RTs for switches (see previous comment).  
  
As pointed out in the manuscript, Minear and Shah (2008) found higher local switch costs for predictive switching and higher global switch costs for random switching. The authors expected and found the reverse pattern in the present study. The results of the Vincentile analyses are furthermore at odds with the results of Huff et al. (2015). First, I have problems with the reasoning for the present hypotheses. If previous research found higher local costs for predictive and higher global costs for random switching, the authors have to convince the reader for expecting the opposite data pattern. From my point of view, it looks like the hypotheses were created post hoc to fit the data pattern. Moreover, the discrepancies of the data pattern with previous literature are just broken down to methodological differences, like the number of trials. From my point of view, such factors cannot explain the complete reverse in a satisfying way.  
  
**Minor Comments**  
- Abbreviations should be used throughout the manuscript once they have been introduced (response times, which are abbreviated on p.3 for the first time but again on p.5)

- Participants: description of the sample is incomplete (mean age, sex)

- Materials: were partial repeats of letter/number pairs (like a repetition of the letter but a switch of the number) excluded for consecutive trials?

- Procedure: why are 59 switch trials and 61 nonswitch trials used in the mixed blocks? This means that no complete pseudorandomization was possible (e.g., of switch/nonswitch and left/right response key)

- Results: were post-error trials removed for RT data analyses? Was the first trial of each block removed?

-Results: Mean values should be given only once, either in the text or in a table, but not both

- Results: Please indicate the complete ANOVA design for each analysis

- Results: Please provide p values for all effects

- Results: How were post-hoc tests corrected?  
  
  
**Reviewer: 3**  
  
The authors report a single experiment that aimed at comparing the magnitude of local and global switch costs among predictable and unpredictable task switching. The main observation consists of larger global switch costs for predictable as compared to unpredictable switching, whereas local switch costs were larger for unpredictable as compared to predictable switching. Apart from effects in mean RTs and error rates, the authors also report a distributional RT analysis.  
  
I found the theoretical rationale of this study relatively weak. On the conceptual level, the authors refer to rather vague notions of working memory and attentional control. Furthermore, they heavily relate to a study of Minear and Shah (2008) that reported data that seem to indicate the opposite pattern of what the authors predicted for their own study without detailing the reasons for this divergence (cf. p. 12, 2nd para). In addition, the rationale of the distributional analysis is mainly stated as conforming to a recent trend in the literature (cf. p. 10). These weaknesses could be acceptable if the reported data were clear cut. However, the reported observations are largely in conflict with previous studies (including one by one of the present authors) and are discussed mainly in terms of some methodological differences (including differences in sample size) without being specific in terms of how these differences in findings may have been brought about. Therefore, I consider the contribution of this paper as too weak to be published in QJEP.  
  
**Specific comments:**  
  
The introduction into task switching research is strongly biased towards the specific paradigm employed by the authors. For example, I doubt that "in most task switching studies" (p. 5, l. 40), pure blocks are employed.  
  
I find the statement that participants "must quickly consider which task-set corresponds to the correct response" (p. 6, l. 45) somewhat strange. I would say that participants must consider which response corresponds to the relevant task set.  
  
p. 8, l. 15: What does "subsequently" mean in this context?  
  
p. 12, l. 33: It is far from clear inasmuch "local switch costs reflect reconfiguration processes". The relative role of reconfiguration and proactive interference for local switch costs is still under debate.  
  
p. 13, ll. 52f: Does "half of the consonants were paired with an odd number, while the remaining half was paired with even numbers" indicate a constant pairing of certain consonants with odd / even numbers? If yes, this would mean that consonant identity was predictive of parity and vice versa.  
  
p. 19, l. 10: It is trivial that "RTs increased across bins".  
  
p. 21, ll. 43f: The observation that local costs were larger in random sequences could simply be due to the fact that participants could use the inter-trial interval for task preparation with predictive alternating runs.