Self-paced preparation for a task switch eliminates attentional inertia but not the performance switch cost

Cai S. Longman, Aureliu Lavric and Stephen Monsell

University of Exeter, Exeter, UK

Abstract: 249 words, Text: 8336 words. 40 References, 3 Figures, 1 Table.

Corresponding author:

Cai S. Longman
Psychology, CLES
University of Exeter
Washington Singer Building
Perry Road
Exeter
EX4 4QG
United Kingdom
tel: +44 (0)1392 724642

fax: +44 (0)1392 724623

email: c.s.longman@exeter.ac.uk

Abstract

The performance overhead associated with changing tasks (the "switch cost") usually diminishes when the task is specified in advance, but is rarely eliminated by preparation. A popular account of the "residual" (asymptotic) switch cost is that it reflects "task-set inertia": carry-over of task-set parameters from the preceding trial(s). New evidence for one component of "task-set inertia" comes from eye-tracking, where the location associated with the previously (but no longer) relevant task is fixated preferentially over other irrelevant locations, even when preparation intervals are generous. Might such limits in overcoming task-set inertia in general, and "attentional inertia" in particular, result from suboptimal 'scheduling' of preparation when the time available is outside one's control? In the present study the stimulus comprised three digits located at the points of an invisible triangle, preceded by a central verbal cue specifying which of three classification tasks to perform, each consistently applied to just one digit location. The digits were presented only when fixation moved away from the cue, thus giving the participant control over preparation time. In contrast to our previous research with experimenter-determined preparation intervals, we found no sign of attentional inertia for the longest preparation intervals. Self-paced preparation reduced but did not eliminate the performance switch cost – leaving a clear residual component in both RT and error rates. That the scheduling of preparation accounts for some, but not all, components of the residual switch cost, challenges existing accounts of the switch cost, even those which distinguish between preparatory and post-stimulus reconfiguration processes.

Key words: task switching, attention, task-set inertia, attentional inertia, eye-tracking.

To perform any cognitive task requires an appropriate "task-set": an organization of processing components and/or tuning of their parameters, including orientation of attention to relevant stimulus attributes, selection of a relevant response modality and set of responses, and activation of appropriate stimulus-response (S-R) rules (cf. Monsell, 2003, 2015). To investigate top-down control of task-set (and the limits of such control) researchers have often employed *task-switching* paradigms (for reviews see: Kiesel et al., 2010; Monsell, 2003, 2015; Vandierendonck, Liefooghe, & Verbruggen, 2010), which contrast performance and other measures, such as brain activity (e.g., Karayanidis et al., 2010; Ruge, Jamadar, Zimmermann, & Karayanidis, 2013; Richter & Yeung, 2014), or eye-movements (Longman, Lavric, & Monsell, 2013, 2016; Longman, Lavric, Munteanu, & Monsell, 2014; Mayr, Kuhns, & Rieter, 2013), for task switches and task repetitions, all other things being equal. This contrast reveals, on switch relative to repeat trials: a longer response time and a higher error rate ("switch cost", Rogers & Monsell, 1995), a posterior positive-polarity preparatory potential in the EEG (e.g., Karayanidis, Coltheart, Michie, & Murphy, 2003; Lavric, Mizon, & Monsell, 2008), one or more negative-polarity post-stimulus EEG potential(s) (e.g., Karayanidis et al., 2003; Astle, Jackson, & Swainson, 2006; Elchlepp, Lavric, Mizon, & Monsell, 2012), greater prefrontal and parietal BOLD fMRI activations (Ruge et al., 2013; Richter & Yeung, 2014), and delays in EEG markers of task-specific processing (Elchlepp, Lavric, & Monsell, 2015) and in shifting the gaze to the relevant stimulus attribute (Longman et al., 2013, 2014; 2016; Mayr et al., 2013).

Although the switch cost reduces, often substantially, with opportunity for preparation (suggesting endogenous task-set reconfiguration, TSR), there tends to remain a non-trivial asymptotic component. Theoretical perspectives vis-a-vis this "residual" component of the switch cost include accounts in terms of task-set reconfiguration processes that cannot occur (Rogers & Monsell, 1995; Rubinstein,

Meyer, & Evans, 2001; Meiran, Chorev, & Sapir, 2000; Meiran, Kessler, & Adi-Japha, 2008), or occasionally "fail to engage" (De Jong, 2000) in advance of stimulus onset, as well as memory-based accounts which attribute the residual switch cost to associative re-activation by the stimulus of (elements) of competing task-sets (Wazsak, Hommel, & Allport, 2003; 2005), or to the greater vulnerability of working memory during a switch to interference from the long-term memory representations of competing task-sets (Mayr, Kuhns, & Hubbard, 2014). However, the account that seems to have received the widest acceptance thus far is that the residual switch cost reflects *task-set inertia*: the carry-over of task-set activation (and/or its inhibition) from the preceding trial(s), which cannot be completely suppressed through preparation (Allport, Styles, & Hsieh, 1994; Yeung & Monsell, 2003).

Unequivocal evidence of task-set inertia is difficult to provide using conventional behavioral measures, so researchers have also employed other dependent variables. With fMRI Yeung, Nystrom, Aronson and Cohen (2006) were able to demonstrate that persisting activation in brain regions associated with the previously (but no longer) relevant task was predictive of the magnitude of the observed RT switch cost; Wylie, Javitt and Foxe (2006) similarly found that the RT cost of switching to a motion- from a color-discrimination task was correlated with activity in neural regions associated with color-processing. In an eye-tracking study we required participants to identify either the face or the letter in a compound face-letter stimulus (Longman et al., 2013). We found that switching induced a tendency to fixate the irrelevant stimulus attribute even when a cue-stimulus interval (CSI) of 800 ms provided ample opportunity for preparation. During a task-switch in the interval when participants made the first fixation on the stimulus (200-400 ms following stimulus onset), they fixated the task-irrelevant attribute for 25% of the time they fixated the relevant attribute, as compared to 17% during a task repetition – a switch-induced increase of 53% in the tendency to

allocate attention to the irrelevant stimulus attribute. Although this finding suggested an attentional component of task-set inertia (which we termed *attentional inertia[1]*), the presence of only one irrelevant stimulus attribute allowed for an alternative interpretation in terms of greater overall distractibility during a switch, and the use of only two preparation intervals made it impossible to ascertain whether the above tendency was indeed 'residual' (asymptotic).

More recently, we examined task-set inertia in a paradigm developed for investigating the dynamics of spatial attention as a component of task-set (Longman et al., 2014). On each trial participants were presented with three digits positioned at the corners of an invisible equilateral triangle (see Figure 1), and asked to perform one of three classifications specified by a small centrally-presented word cue ("ODD?", "EVEN?", "LOW?", "HIGH?", "INNER?", or "OUTER?"; see Methods for the number categories in these three classification tasks) that had to be fixated at the start of the trial. Each classification was consistently associated with a location, so that the cue specified both the task and the relevant location. The use of three tasks allowed comparison, on switch trials, between the probabilities of fixating the location associated with the task relevant on the preceding trial and of fixating the location associated with the other irrelevant task; preferential orientating towards the former location would rule out a mere distractibility account. We also used four CSIs (120 ms, 620 ms, 1020 ms and 1420 ms) to determine whether any attentional inertia was indeed 'residual' (asymptotic). In a control (location switching) condition, participants switched locations while performing the same task throughout the experiment. At all CSIs, task switching induced both a delay in fixating the task-relevant region of the stimulus and a tendency on switch trials to orient preferentially towards the irrelevant

---

[1] The term "attentional inertia" has previously been used by Kirkham, Cruess and Diamond (2003) to explain young children's difficulty when shifting between sorting cards by colour or by shape. The time-scale of the persistence of the children's "attentional inertia" seems rather different from the transient one-trial inertia we detect in adults; it remains an open question how they are related.

digit location which was the relevant location on the previous trial. This attentional inertia was reduced in conditions with longer preparation intervals but a substantial bias towards fixating the previously (but no longer) relevant region was still reliable at the two longest CSIs, and the pattern of means in these CSIs showed that it had reached an asymptotic minimum. There were no comparable delays or attentional inertia in the (location switching) control condition.

In the standard task-cuing experiment, the experimenter controls the cue-stimulus interval, whether it varies from trial to trial, or is held constant over a block (as in Longman et al.'s studies). This raises the question of whether conforming to an experimenter-imposed deadline militates against optimal scheduling of reconfiguration of task-set components. Consider attentional inertia as documented by Longman et al. (2014) in the longest CSI conditions: the majority of the fixations on the previously (but no longer) relevant stimulus element were made early in the preparation interval, before the mean latency of the appropriately targeted fixations. This suggests that a time limit on preparation may have encouraged rapid deployment of spatial attention (before attentional parameters had been fully reconfigured), perhaps in order to allow other TSR processes to also occur during the preparation interval. This may, at least in part, explain the vulnerability to the carryover of the most recent attentional settings. The tendency to orient promptly was unlikely to be due to insufficient preparation time – indeed our longest CSI of 1420 ms was rather generous (in task-switching experiments asymptotic preparation is typically achieved well within one second). Instead, it was more likely due to the difficulty in optimizing the timing of the spatial orienting as part of readying the task-set for the arrival of the stimulus. The above logic can be extended to the rest of the task-set. When there is an externally-imposed deadline, the effort to "synchronize" one's state of optimal readiness with the onset of the stimulus could be hampered by factors such as trial-to-trial variability in the duration of TSR processes

and the difficulty of estimating the time available for preparation, especially when the preparation interval changes. Subjects may therefore set suboptimal deadlines for TSR components/stages.

The current experiment tested the above conjectures. The most obvious way to determine whether suboptimal 'scheduling' of TSR is responsible for the residual switch cost is to allow the participant to control the timing of the target stimulus onset. Eye-tracking makes it possible to grant the participant control over the duration of the preparation interval without requiring any extra behavior (such as making an extra key press). The cues and stimuli were identical to those used by Longman et al. (2014; see Figure 1), but the participant was able to initiate stimulus onset simply by shifting their gaze from the cue to any of the three critical stimulus regions – a *gaze-contingent display* paradigm. Although using gaze to put the CSI under the participant's control is an innovation in task switching, our paradigm bears some resemblance to that used by Arrington, Logan and Schneider (2007), who required participants to make discriminative vocal or key press responses to the task cue in order to allow a separation of task switch costs into cue encoding costs and stimulus processing costs. However, our design differs from theirs in that we did not require the subject to make any discriminative response to the cue that they would not otherwise make, or categorize the cue. Instead, we simply instructed the participants to prepare for the upcoming task before shifting their gaze away from the cue, thus making explicit the relationship between the latency of the first eye-movement away from the cue and the duration of the preparation interval. Consequently, although the time between cue onset and the gaze leaving the cue in the present paradigm includes the cue encoding time, it will also include the duration of task preparation processes per se, whilst not adding any additional processing required to explicitly signal the readiness for the imperative stimulus.

Our paradigm is also reminiscent of the "strategy cuing paradigm" developed by Dixon and colleagues (Dixon, 1981; Dixon & Just, 1986): in a choice reaction task the stimulus-response mapping changed from trial to trial; a cue presented in advance specified the stimulus-response mapping. Subjects could take as long as they wished to inspect the cue, after which they could initiate the trial by pressing a foot pedal. Dixon et al. used the time elapsed between cue onset and trial initiation as a measure of the time it took to prepare a response strategy. In the task-switching domain this paradigm has been adapted by Hübner, Futterer and Steinhauser (2001) who, on every trial, presented the participants with a visual cue that specified one of two tasks and asked them to press a key (with the hand that was not subsequently used for responding) when they were ready for the target stimulus. They reported RT switch costs of 20-50 ms (Experiment 3)[2], which seemed to suggest that self-paced preparation did not eliminate the switch cost. However, in their design the task cue changed on switch trials and it was repeated on task repeat trials – hence one cannot disentangle the 'true' task switch cost from the facilitation arising from cue repetition. (The latter has subsequently been found to be substantial, e.g., Logan & Bundesen, 2003; Monsell & Mizon, 2006). In the present paradigm, the cue changed on every trial (even when the task was repeated, e.g., "ODD?" or "EVEN?" for the parity task), enabling us to unconfound the effects of task and cue change. Making the onset of the target stimulus contingent on gaze position may result in an extra demand – that participants ensure that their gaze does not leave the cue until they are prepared for the stimulus. However, the data from our previous eye-tracking experiment (Longman et al., 2014), showed that even in the absence of this requirement, participants spent a relatively long time fixating the cue (400-600 ms), and that there was no discernible tendency for very early eye-movements away from the cue

---

[2] What the authors referred to as "shift cost" was the comparison of mixed-task blocks (containing task switches) with single-task blocks. However, this measure conflates the task switch cost with the "mixing cost" – the difference between task repetitions in mixed task blocks and single-task blocks. But the authors also reported the more widely accepted measure of trial-to-trial switch cost as the difference between switches and repetitions within mixed-task blocks – it is this analysis we refer to.

(<300 ms from cue onset) – hence we assume there was no obvious need to suppress such a tendency. Furthermore, the use of gaze-contingent displays makes it possible to avoid the extra manual (or foot) response required in Dixon and colleagues' and Hübner and colleagues' studies. Finally, by recording eye movements from cue onset, in addition to the timing and the accuracy of responses to the stimulus, it was possible to establish whether allowing the subject to control the duration of the CSI eliminates both the attentional inertia we previously reported and the residual switch cost observed in most task-switching studies.

Method

Participants

Eighteen participants (10 female, mean age = 21.54) were either paid £10 (£5/hr) or awarded partial course credits plus a performance-related bonus of up to £4.50 for their participation.

Tasks, stimuli and procedure

The experiment was run using Experiment Builder (SR Research, Ottawa, Canada) on a flat-screen (zero curvature) 17" CRT monitor placed ~57 cm from the participants' eyes. Each trial started with a blank screen followed by a centrally presented black fixation cross (subtending 0.4°). Allowing the subject to control the preparation interval means that the CSI and the response-stimulus interval (RSI) must necessarily be confounded. Independent manipulation of CSI and RSI has yielded separable effects on the switch cost, typically interpreted as reflecting endogenous TSR vs. passive decay of task-set, respectively (Meiran, 1996). We therefore added a design parameter to enable analyses that could unconfound the two kinds of effect – we presented the fixation cross at one of three response-fixation intervals (RFIs; 700 ms, 1200 ms, 1700 ms; randomized over trials, but equiprobable for each task, transition

9

type and level of each element within the stimulus). After the fixation cross had been visible for 400 ms (minimum) and then fixated for a further 100 ms it was replaced with a task-cue ("ODD?", "EVEN?", "LOW?", "HIGH?", "INNER?", "OUTER?" subtending up to 0.8° horizontally, 0.3° vertically) accompanied by three light blue dots (0.3° in diameter) at the locations where the digits would eventually appear (see Figure 1). Trials on which the interval between the onset of the fixation cross and the onset of the cue exceeded 1000 ms were assumed to reflect problems either achieving or detecting fixation and were omitted from all analyses (1.07%). The cue changed on every trial, thereby unconfounding task switching and cue switching (Monsell & Mizon, 2006). Participants were instructed that the cue would remain visible as long as it was fixated and that the stimulus would appear as soon as they moved their eyes towards one of the stimulus elements. This was implemented by presenting the stimulus as soon as the gaze crossed the boundary of one of the invisible square (side = 2°) regions centered on each of the three stimulus elements. Participants were explicitly informed to take as long as they needed to prepare for the upcoming stimulus, which would appear as soon as they moved their eyes away from the cue. Consequently, they controlled the time available for preparation. The stimulus remained on screen until a response was made by pressing one of the keys ('c' or 'm') on a standard QWERTY keyboard. If an incorrect response was made "ERROR" was displayed for 1000 ms before the start of the next trial.

The stimulus comprised three digits, selected from the set 2-9 (each subtending 0.4° horizontally, 0.5° vertically), displayed at the points of an invisible equilateral triangle, 5° from each other and 2.7° from the central location of the cue. The cue specified the classification task to apply: odd vs. even; low (≤5) vs. high (≥6); "inner" (4, 5, 6, 7) vs. "outer" (2, 3, 8, 9); with each task mapped to a single location throughout the experiment (with allocation of task to location counterbalanced over participants).

Stimuli were generated by first defining the level (e.g., odd, high) of each digit and then randomly selecting a digit from the appropriate set with two constraints: a digit could appear only once in a stimulus and there could be no exact repetition from the previous trial of the whole stimulus (the same three digits in the same locations). Stimuli were also controlled to ensure that the combinations of categories of the three digits (e.g., odd, high, outer) were equiprobable for each transition type, task and RFI.

The tasks were equiprobable, so the task switch:repeat ratio was 2:1. We assume TSR, especially its preparatory component, to be endogenous, hence sensitive to motivational factors (cf. Nieuwenhuis & Monsell, 2002). Thus, to encourage endogenous preparation, we used the following monetary incentive scheme. The mean RT, number of errors and a composite performance score [(mean RT in ms/10) + (errors*5)] were displayed at the end of each block. Bonus payments were awarded for improvement relative to the average performance on previous blocks. We have employed this incentive procedure extensively in task-switching studies, some of which used eye-tracking (Longman et al., 2013; 2014; 2016), and some of which did not (Lavric et al., 2008; Elchlepp et al., 2012; 2015) – these studies found robust behavioral, eye-movement and electrophysiological indices of effective preparation.

Before the eye-tracking session, participants practiced each task in an 8-trial block. This was followed by a 74-trial practice task-switching block in which only one digit was presented centrally on each trial. Location-task mappings and the gaze-contingent timing of the cue-stimulus interval were then introduced in two further practice blocks of 74 trials each, identical to the fifteen 74-trial experimental blocks that followed resulting in 1100 experimental trials. The whole session, including the practice and experimental blocks and setting up the eye-tracker, lasted two hours.

Eye-tracking

An EyeLink 1000 eye-tracker (SR Research, Ottawa, Canada) recorded the movements of the right eye from the onset of the blank screen until the response with a sampling rate of 500 Hz; it was calibrated before each block. Four square regions (side = 2°) were delineated as containing 'relevant' information for the three digits in the stimulus and the cue. The proportion of all fixations to land in these regions was 91.19% on both switch and repeat trials. The first two trials in each block, trials with RT>2500 ms (1.14%) and trials following an error were omitted from all analyses. Error trials were also omitted from RT and eye-tracking analyses.

Only fixations initiated after cue onset were analyzed. The total dwell-time on each of the above-mentioned spatial regions of the stimulus within a defined time-window (the summed duration of all fixations on the respective region within the respective time-window) was submitted to statistical analysis.

In ANOVAs, the reported significance values were Huyhn-Feldt-corrected for sphericity violations (but *dfs* are reported uncorrected).

Results

Cue-stimulus interval (CSI)

Participants spent more time fixating the cue before initiating an eye-movement away from it on switch (741 ms) than on repeat trials (674 ms), see the inferential statistics for this comparison below. In order to assess the extent to which the time taken to prepare influenced performance, trials were grouped into quartiles, separately for each participant's switch and repeat trials, according to the duration of this self-paced CSI. The mean CSIs for the quartiles, averaged over switch and repeat conditions (and the standard deviation of the mean over participants), were: 466 (62) ms, 558 (81) ms, 671 (115) ms, 1134 (341) ms. We also computed, as a measure of the range of CSIs for

each quartile, the standard deviation over individual trials for each participant; the averages of those over participants were: 48 ms, 48 ms, 83 ms, 519 ms. All performance and eye-tracking data were analyzed including these quartiles as levels of the factor CSI.

Figure 2 shows the mean CSIs, RTs and errors averaged across all three tasks and RFIs. The CSI durations were submitted to an ANOVA with the factors Switch, CSI quartile, Task and RFI; this found a reliable main effect of Switch, $F(1,17)=33.62$, $MSE=44339.10$, $p<.001$, indicating that the cue was fixated for 68 ± 12^3 ms longer on switch than repeat trials. The switch-repeat difference in CSI duration increased from the shortest CSI quartile to the longest (Switch by CSI interaction, $F(3,51)=9.95$, $MSE=46202.80$, $p=.005$). This difference was reliable for all quartiles (main effect of Switch at CSI 1: 32 ± 5 ms, $F(1,17)=47.83$, $MSE=1786.21$, $p<.001$; CSI 2: 45 ± 6 ms, $F(1,17)=56.02$, $MSE=2910.32$, $p<.001$; CSI 3: 61 ± 10 ms, $F(1,17)=38.04$, $MSE=7820.60$, $p<.001$; CSI 4: 134 ± 32 ms, $F(1,17)=17.64$, $MSE=81684.93$, $p=.001$), suggesting that switching was associated with a significant delay in orienting task-relevant spatial attention across the whole range of CSIs. There were no other reliable interactions involving the factor Switch in the analysis of CSI durations.

RT and errors

An ANOVA with the factors Switch, CSI, Task and RFI on the RTs (see Figure 2) found the RT to increase with CSI, ($F(3,51)=35.71$, $MSE=34229.87$, $p<.001$). There was a reliable overall switch cost of 82 ± 14 ms, $F(1,17)=32.80$, $MSE=66191.01$, $p<.001$, which reduced with the self-paced preparation time from 112 ± 16 ms at the shortest CSI quartile to 45 ± 17 ms at the longest, $F(3,51)=8.34$, $MSE=7321.65$, $p<.001$ (Switch by CSI interaction). The 'residual' switch cost was statistically significant in the longest CSI quartile, $F(1,17)=6.81$, $MSE=24573.98$, $p=.018$; all participants but one had a longer mean RT on switch than repeat trials.

---

[3] Standard error of the contrast.

Could the effects of CSI reflect not the time available for preparation, but the response-stimulus interval (RSI; time available for passive dissipation after the previous trial)? As already mentioned (see Method) average RSI was necessarily longer at longer CSIs. To determine whether there was a detectable effect of RSI (which varied with RFI, CSI and, to a much lesser extent, the fixation-cue interval) we selected a subset of trials for which the RSI was within one standard deviation of the median RSI for the condition (collapsed across all three tasks). This resulted in a systematic variation in RSI with RFI (mean RSIs in order of increasing RFI were: 1858 ms, 2374 ms, 2895 ms), yet, for the first three quartiles, the CSI was comparable for each RFI within a CSI quartile. The mean CSIs in order of increasing RFI for CSI quartile 1 were: 472 ms, 468 ms, 473 ms; CSI quartile 2: 552 ms, 555 ms, 560 ms; CSI quartile 3: 650 ms, 657 ms, 676 ms. CSI quartile 4 was not included in the analyses which examined the effect of RFI within each CSI because the CSI duration increased substantially and monotonically as a function of RFI in this quartile (mean CSIs in order of increasing RFI: 900 ms, 954 ms, 1003 ms) thereby confounding the time available for passive dissipation (whose effects we aimed to examine here) with the preparation interval. An ANOVA with the factors RFI and Switch on the RTs from each of the three shortest CSI quartiles (only including those trials selected as described above) found that the effects of RFI per se on overall performance were modest: the switch cost in order of increasing RFIs for CSI 1 was: 100 ± 18 ms, 140 ± 27 ms, 118 ± 25 ms; CSI 2: 69 ± 33 ms, 102 ± 23 ms, 96 ± 34 ms; CSI 3: 53 ± 29 ms, 80 ± 14 ms, 86 ± 38 ms. Neither the main effect of RFI nor its interaction with Switch approached significance for any of the CSIs (all Fs<2). The results from this analysis suggest that the reduction in switch cost with increasing CSI cannot be explained by passive dissipation and is therefore a consequence of endogenous preparation – indeed, for all three CSI quartiles included in the analysis, the switch cost increased numerically from the shortest RFI to the longest.

A reliable switch cost was also found in the errors, $F(1,17)=5.13$, $MSE=66.43$, $p=.037$, with participants making $1.0 \pm 0.5\%$ more errors on switch than on repeat trials. This effect was not significantly modulated by preparation, $F(3,51)<1$, but there was a reliable interaction between Switch and Task, $F(2,34)=4.88$, $MSE=17.00$, $p=.014$. The error switch cost was $1.8 \pm 0.6\%$ in the odd/even task, $F(1,17)=8.89$, $MSE=39.21$, $p=.008$, $1.2 \pm 0.5\%$ in the low/high task, $F(1,17)=5.04$, $MSE=31.21$, $p=.038$, and $0.07 \pm 0.5\%$ in the inner/outer task, although the latter was not significant, $F<1$.

Analyses with CSI as a continuous variable. To ensure that the reduction in the RT switch cost with CSI reported above is not dependent on our choice of quantile for partitioning the distribution of self-paced CSI (or indeed any choice of quantile), we also tested the effect of CSI on the switch cost by treating CSI as a continuous variable in a Linear Mixed Effects (LME) analysis on single-trial RTs, using the lmer.test package in R. We tested five models, all of which included trial order as a fixed effect and participant as a random effect whose intercept was allowed to vary. Four models included the key interaction, CSI x Switch, along with the main effect of Switch and the Switch x RFI interaction, plus: no other terms (model 1); the Task x Switch interaction (model 2); the Task x Switch x CSI interaction (model 3); the Task x Switch x RFI interaction (model 4). In all four models (see Supplementary Materials for the R syntax and ANOVA output) CSI had a significant effect on the RT switch cost ($p<.001$ for the CSI x Switch interaction). To ascertain that the CSI x Switch interaction was not specific only to a subset of models, we also tested the fifth model, which contained all the main effects in the design (CSI, Switch, RFI and Task) and all the interactions among them. In this unparsimonious but comprehensive model the CSI x Switch interaction was again significant ($p<.05$). The RFI x Switch interaction was not significant in any of the five models tested. Thus, when we treated CSI as a continuous

variable, we confirmed the statistically robust effect of self-paced CSI (and no statistically detectable effect of RFI) on the switch cost.

Eye-tracking

      Delay in fixating the relevant region. Figure 3 shows the average dwell time in successive 20 ms bins for the three stimulus regions described above (see Method) as a function of switch/repeat and CSI[4]. Dwell time is the amount of time spent fixating each of the stimulus regions per 20 ms time-bin per trial. Our dwell-time measure is analogous to the probability of fixating a given region at a given moment in time (the probability can be calculated by multiplying the value presented in Figure 3 for each 20 ms bin by 0.05). We estimated the apparent delay in appropriate orienting of attention on switch trials relative to repeats by first selecting a 200 ms window[5] on the repeat curve starting from the 20 ms bin half way between cue onset and the first bin in which dwell time (averaged over participants) exceeded 95% of the maximum of the entire dwell-time function (see Figure 3, left panels, for the window boundaries). For the dwell-time (y axis) value at every 2 ms point on the repeat curve we estimated the delay in the switch curve reaching this value using the same non-linear curve fitting procedure as in our earlier studies (Longman et al., 2014, 2016)[6]. The estimated switch-related delays in order of increasing CSI quartile were: 131 ± 23 ms, 85 ± 13 ms, 63 ± 13 ms, 186 ± 73 ms). Four one-sample t-tests (one for each CSI) found that the estimated delay was reliable for all CSIs (CSI 1: $t(17)=5.74$, $p<.001$; CSI 2: $t(17)=6.51$, $p<.001$; CSI 3: $t(17)=4.65$, $p<.001$; CSI 4: $t(17)=2.53$, $p=.021$). A follow up ANOVA with the factor

---

[4] The apparent discrepancy between the mean CSI duration (initial fixation landing time) and the dwell time curves in Figure 3 (early fixations at the longest CSI are prior to the mean latency of the first fixation at the shortest CSI) is due to considerable variability between participants/conditions in the self-paced CSI durations.

[5] Because the portion of the dwell time curve with maximal rise in CSI quartile 4 was considerably longer than in the other CSIs, we used a 400 ms window there.

[6] In a procedure analogous to high-order spline interpolation, we fitted a 10-order polynomial (Polyfit function in Matlab; Mathworks, Natick, MA) to the switch curves of every subject. The real (noncomplex) polynomial root corresponding to a y-axis value on the repeat curve is an estimate of the time point at which the switch curve reached this value. Averaging these estimates within the 200-ms time-window defined above provided the mean temporal separation between the switch and repeat curves.

CSI found that the differences in the delay in appropriate attention orientation with CSI quartile did not approach significance, $F(3,51)=2.19$, $MSE=65526.14$, $p=.154$.

Dwell time on the irrelevant regions. The dwell time on the currently irrelevant regions (Figure 3, right panel) was analyzed in two contiguous 500 ms windows starting 200 ms after cue onset; any one of these analysis windows would fail to capture the interval of maximal dwell-time on these regions for all CSI quartiles, whereas a wider window could reduce sensitivity by 'diluting' potential effects by including portions with negligible dwell-time. ANOVAs with the factors Switch, CSI, Task and RFI were run on the dwell time in the early and late windows separately. Both analyses found a greater tendency to fixate the irrelevant regions (more dwell time) on switch trials than on repeat trials [early window: 5.1 ± 1.1 ms, $F(1,17)=18.90$, $MSE=441.44$, $p<.001$; late window: 6.9 ± 1.5 ms, $F(1,17)=20.11$, $MSE=760.13$, $p<.001$]. This effect was reduced at longer preparation intervals in both windows [early: $F(3,51)=10.10$, $MSE=257.09$, $p<.001$; late: $F(3,51)=4.22$, $MSE=309.62$, $p=.017$], but remained reliable (or approached significance) in all but the longest CSI quartile [early window: main effect of Switch in CSI 1: 12.5 ± 2.4 ms, $F(1,17)=26.49$, $MSE=476.38$, $p<.001$; CSI 2: 5.5 ± 2.5 ms, $F(1,17)=4.90$, $MSE=506.49$, $p=.041$; CSI 3: 2.6 ± 1.4 ms, $F(1,17)=3.32$, $MSE=169.15$, $p=.086$; CSI 4: -0.4 ± 0.4 ms, $F(1,17)<1$; late window: main effect of Switch in CSI 1: 7.5 ± 2.2 ms, $F(1,17)=11.83$, $MSE=381.43$, $p=.003$; CSI 2: 9.3 ± 2.8 ms, $F(1,17)=10.72$, $MSE=649.21$, $p=.004$; CSI 3: 9.1 ± 2.0 ms, $F(1,17)=21.51$, $MSE=309.94$, $p<.001$; CSI 4: 1.7 ± 1.4 ms, $F(1,17)=1.54$, n.s.]. There was also a reliable interaction between Switch, CSI and Task in the early window, $F(6,102)=4.43$, $MSE=294.46$, $p=.003$, which approached significance in the late window, $F(6,102)=2.02$, $MSE=359.23$, $p=.075$. Follow-up analyses by Task found a systematic reduction in the tendency to fixate the irrelevant regions on switch trials

relative to repeats with increasing CSI (Switch by CSI interaction) only for the inner/outer task in the early window, $F(3,51)=11.85$, $MSE=422.92$, $p<.001$.

Attentional inertia. For the crucial analysis of attentional inertia (see Introduction), we submitted to an ANOVA the dwell time in the currently irrelevant regions on switch trials with the factors Previous Relevance (previously relevant, previously irrelevant), CSI, Task and RFI, for the same time-windows as above. Figure 3 (right panels) and Table 1 (which presents the data by both CSI and RFI) shows a tendency to fixate the previously relevant region for the two shorter CSIs, disappearing at the longer CSIs. In the early window this attentional inertia effect reduced as a function of CSI, $F(3,51)=7.19$, $MSE=3050.00$, $p=.005$. The effect was significant for the shortest CSI quartile, $F(1,17)=7.93$, $MSE=6318.45$, $p=.012$, but not for the second ($F<1$). For the two longest CSI quartiles the effect was even non-significantly reversed [CSI 3, $F(1,17)=1.68$, n.s.; CSI 4, $F<1$], indicating that attentional inertia had been eliminated at the longer CSIs.

The ANOVA that examined attentional inertia in the late window found an overall tendency to fixate the previously relevant region *less* than the other irrelevant region, $F(1,17)=7.82$, $MSE=1228.28$, $p=.012$. The Previous Relevance by CSI interaction approached significance, $F(3,51)=2.48$, $MSE=359.25$, $p=.088$, but inspection of Table 1 shows no systematic modulation of the effect by CSI. Follow up ANOVAs demonstrated that the tendency to fixate the previously relevant region less than the other irrelevant region was reliable only for CSI quartiles 2 and 3 [main effect of Previous Relevance in CSI 1, $F(1,17)=2.78$, $MSE=265.27$, $p=.114$; CSI 2, $F(1,17)=6.39$, $MSE=974.12$, $p=.022$; CSI 3, $F(1,17)=6.27$, $MSE=652.19$, $p=.023$; CSI 4, $F(1,17)=3.91$, $MSE=177.29$, $p=.064$].

Active reduction or passive dissipation of inertia? Neither the main effect of RFI, $F(2,34)<1$, nor its interaction with Previous Relevance, $F(2,34)<1$, were reliable in

the omnibus ANOVA for the early window. To ascertain whether the above-mentioned reduction and elimination of our index of attentional inertia with increasing CSI reflected the duration of the preparation interval, rather than passive dissipation of the previous task-set parameters due to a correlated increase in the response-stimulus interval (RSI), follow up analyses were performed for the three shortest CSIs with the factors RFI and Previous Relevance (using only the subset of trials selected in the same way as for the corresponding RT analysis above, where we explained why for these trials the jitter in RFI is the primary source of variance in RSI, and why CSI quartile 4 was excluded from this analysis). The main effect of RFI approached significance for the shortest CSI, but was not reliable for the longer CSI quartiles [main effect of RFI at CSI 1, $F(2,34)=3.13$, $MSE=243.94$, $p=.061$; CSI 2, $F(2,34)=1.92$, $MSE=228.05$, $p=.168$; CSI 3, $F(2,34)=2.42$, $MSE=73.39$, $p=.104$]. More importantly, the RFI by Previous Relevance interaction did not approach significance for any CSI (all Fs<1). In Table 1 the reduction in the preference for the previously relevant region over RFIs is the trend over rows for each column; the preference is 5.3 ms in CSI 1 and <1 ms in the remaining CSIs. The same reduction over CSIs is the trend over columns for each row: ~22-28 ms over different RFIs. Thus, the numerical reduction of the preference for the previously relevant region over the other irrelevant region with increasing RFI was trivially small relative to the effect of increasing CSI. Hence the effect of CSI can be attributed largely to the duration of the preparation interval, not to the time available for passive dissipation of attentional settings.

Analyses with CSI as a continuous variable. As for the RTs, to confirm that these outcomes are not dependent on the choice of the CSI quantile, we ran (for both the early and late windows) analyses treating CSI as a continuous variable. LME analyses were performed on the index of attentional inertia – the previously relevant-other irrelevant difference – computed for individual switch trials. As for the RTs, we tested

the most comprehensive model including CSI, RFI and Task (and all the interactions),

plus four restricted models that included the factors CSI and RFI with or without the

main effect of Task or its interactions with CSI and RFI (see Supplementary Materials).

Consistent with the distribution-based analyses above, in the early window CSI, but not

RFI, had a significant effect on attentional inertia (main effect of CSI $p<.01$ in all

models). In the late window, there were no significant effects of either CSI or RFI.

Discussion

We investigated the robustness of attentional inertia in a task-cuing experiment,

with three tasks consistently associated with different locations in the stimulus array,

and the duration of the preparation interval controlled by the participant. Participants

triggered the stimulus onset by shifting their gaze from the central cue to one of the

elements of the stimulus array, with no limit imposed on the time available for

preparation. This use of gaze-contingent displays gave the participant control over

preparation time but, in one respect at least, the behavioral effects observed were similar

to those commonly reported in task-cuing experiments: there was a substantial RT cost

of a task switch that reduced with, but was not eliminated by, a longer preparation

interval.

The effect of preparation interval was assessed by dividing the trials into

quartiles based on the duration of the self-paced CSI. Because the time taken by the

participant for preparation was necessarily confounded with the time available for

passive dissipation, we introduced extra variability into the RSI by jittering the intertrial

interval preceding the onset of the fixation cross (the RFI). The analysis of the effect of

RFI and its interaction with switching in each CSI quartile, and an LME analysis

treating CSI as a continuous variable, both clearly indicated that the reduction in switch

cost with increasing CSI reflected time available for active preparation, not passive dissipation of the previous task-set (or any other effect of an increase in RSI per se).

One aspect of the current data concerning the relationship between CSI and performance is somewhat unusual in comparison to the pattern seen in the standard task-cuing experiments: overall mean RT was longer with increasing CSI and the reduction in the RT switch cost with CSI seemed to arise primarily from the increase in RT on repeat trials at longer CSIs (see Figure 2), whereas what one typically observes in task-switching experiments is an overall reduction in RT with CSI and a greater reduction for switches than for repeats (e.g., Monsell & Mizon, 2006). We speculate that this is due to fluctuation in processing efficiency over trials: participants take longer to prepare when they experience a dip in processing efficiency, so that general performance worsens with a longer CSI. An overlap of this general trend (common to both switches and repeats) with a counteracting trend of switch-specific preparation can explain the observed trend of a smaller increase in RTs on switch trials than on repeat trials with an increasing CSI (and the resulting reduction in switch cost at the longer CSIs). This dissociation between a relative improvement in switch-specific preparation with increasing CSI, and a decline in overall performance, provides some support for the attribution of the two effects of CSI in standard task-cuing experiments to different sources: TSR on the one hand, and on the other the generic preparation responsible for the classic effect of foreperiod on RT.

In our previous research (Longman et al., 2014, Experiment 2, see Introduction) we found that task switching resulted in preferential fixation of the previously relevant stimulus location on switch trials– attentional inertia – even when the opportunity for preparation was ample. In Longman et al.'s (2014) condition with a CSI of 1020 ms the previously (but no longer) relevant stimulus element was fixated on average for an extra 7.8 ms per trial (an extra 82%), relative to the irrelevant stimulus element that was also

irrelevant on the previous trial; this effect was still 5.3 ms (62%) when the CSI was 1420 ms. In the present experiment the attentional inertia effect was statistically significant only in the early window and was numerically altogether absent (even reversed) for the two quartiles comprising the longest CSIs (see Figure 3 and Table 1), even though the mean of the slowest (4th) quartile of self-determined CSIs was shorter than the longest experimenter-determined CSI in our previous experiment. Indeed, for the 4th CSI quartile in the early analysis window, the previously relevant region was fixated 0.02 ms per trial *less* than the other irrelevant region. The 95% confidence interval around this very small mean effect was -1.16 ms to 1.12 ms, and it is clear that the above-mentioned attentional inertia effects of 5-8 ms in Longman et al.'s (2014) long CSIs fall well outside this interval. These results suggest that the tendency to fixate the previously (but no longer) relevant region during a task-switch – an index of task-dependent carryover of spatial attentional settings – can be eliminated when preparation is not constrained by an experimenter-defined deadline.

It is worth noting that the mean CSI in the longest self-determined CSI quartile in the present study roughly equaled the mean of the two longest CSIs (1020 ms and 1420 ms) in Longman et al. (2014, Experiment 2) for which there was clear attentional inertia. We compared the "residual" component of attentional inertia in the two experiments with an independent samples t-test contrasting the estimate of attentional inertia (the previously relevant-other irrelevant dwell-time difference on switch trials) in the current experiment vs. Longman et al.'s (2014) Experiment 2. The difference in the residual attentional inertia was statistically significant, $t(40) = 2.17$, $p=.036$. Comparing the reduction in inertia with an increasing CSI is less straightforward, because of the large difference between the mean CSI in the 1st quartile in the present experiment (466 ms) and the shortest CSI of 120 ms in Longman et al. (2014). We therefore computed a basic average slope estimate of the reduction in inertia with preparation, by taking the

difference in inertia between the longest and shortest CSIs and dividing it by the difference in CSIs. An independent samples t-test comparing the slopes in the two experiments found the reduction in inertia to be significantly steeper in the present experiment than in Longman et al.'s (2014) Experiment 2, $t(40) = 2.19$, $p=.034$.

Thus, similar preparation intervals resulted in markedly different fixation patterns depending on whether the preparation interval was determined by the participant (the present study) or by the experimenter (Longman et al., 2014). Presumably, when the duration of the preparation interval is externally determined (even if the interval is long) participants sometimes reorient attention before the appropriate settings have been updated for the current task-goal (at least on a proportion of trials), see Introduction. The opportunity to self-pace preparation is likely to strongly reduce this tendency to orient "prematurely", allowing for more time to counteract "attentional inertia". Also, an externally imposed long CSI coupled with our relatively poor ability to estimate temporal intervals, may encourage the participant to prepare early, but maintaining an optimal state of preparation for as much as a second may be difficult. Under self-paced conditions, the state of preparation can be used as soon as it is achieved. More generally, the benefit of self-paced preparation is that it allows the stimulus onset to be determined by the participants' meta-awareness of their readiness. Assuming this meta-awareness is at least moderately accurate, self-pacing preparation may result in better synchronization of preparation with stimulus onset. On the trials when TSR, or earlier processing stages (e.g., cue encoding) happen to be slow, preparation can be allowed to take longer, whereas on trials when pre-TSR and TSR processes are effective, stimulus onset can be initiated promptly, before the newly-reconfigured task-set parameters start to decay (see Logan & Gordon, 2001, for a related discussion of the time-course of preparatory executive control processes as a source of the switch cost).

We have recently identified another circumstance in which attentional inertia was eliminated – when instead of cues transparent with regard to the required categorization (as in the present experiment) we used explicit location cues: arrows and the spatial adverbs "TOP", "LEFT", and "RIGHT" (Longman et al., 2016). However, other measures provided very different outcomes in that study compared to the present investigation (and to Longman et al., 2014). First, the use of explicit location cues eliminated not only attentional inertia, but also the delay in fixating the task-relevant region (here and in Longman et al., 2014, the delay was present for all the preparation intervals). Second, the switch cost (and its residual component) increased substantially with explicit location cues relative to Longman et al.'s (2014) experiments, whereas here the residual cost is somewhat smaller than in the latter experiments (see below). One must therefore conclude that different factors led to the elimination of attentional inertia in Longman et al.'s (2016) study and the present study. In the former, it was the de-coupling of the reconfiguration of spatial attention from other task-set reconfiguration processes due to (quasi)automatic orienting triggered by the explicit location cues. In the present study, it is the more optimal scheduling of the anticipatory reconfiguration processes, including eye-movements to the relevant location on the screen, in the absence of an externally-imposed deadline.

Although there was no sign of attentional inertia in the long CSI trials, there was a greater tendency to fixate either task-irrelevant digit (irrespective of its relevance on the preceding trial) on switch relative to repeat trials. This tendency may reflect weaker (and noisier) activation of attentional settings on switch trials or increased distractibility or some combination of the two. There was little effect of RSI (as assessed by analyzing RFI orthogonally to CSI) on fixations in the irrelevant regions suggesting that passive dissipation of attentional settings (or any other effect of an increasing RSI) cannot explain away the above-mentioned effects of CSI on attentional dynamics.

The analysis of dwell-time on the task-relevant region, displayed in the left panel of Figure 3, found that a task-switch delayed orientation to this region – the delay (whose variation over CSIs, 63-186 ms, was non-significant) was statistically significant for each CSI. Thus, as in our work with fixed CSIs (Longman et al., 2014), the delay in appropriate orienting of attention (indexing the time taken for reconfiguration of task-relevant attentional settings) seems to be substantial whatever the duration of the preparation interval. This indicates that preparation for a task switch remains a time-consuming process whether it is self-paced or constrained by the experimenter-imposed CSI.

The evidence of this study indicates that on about 50% of the trials (the two longest CSI quartiles) participants took full advantage of self-paced preparation, making it possible on those trials to overcome the task-switch-related inertia in spatial attention observed with a fixed preparation interval. But does it also eliminate the residual switch cost in RT and errors? The answer is clearly "no": there was still a substantial RT cost at the two longer CSIs where the attentional inertia effect was eliminated. The RT switch cost reduced by nearly two thirds from the shortest to the longest CSI, and this reduction (from 112 ± 16 ms when CSI = 466 ms to 45 ± 17 ms when CSI = 1134 ms) is proportionately larger than the reduction we observed previously with fixed CSIs (from 124 ± 16 ms when CSI = 120 ms to 74 ± 11 ms when CSI = 1420 ms; Longman et al., 2014). Using the same procedure as for the dwell-time index of attentional inertia (see above), we computed an average slope estimate of the RISC by subtracting the switch cost in the 4th CSI quartile from the switch cost in the 1st CSI quartile and dividing this by the difference in the mean CSI for the two quartiles; the same calculation was also done for the switch costs for the shortest and longest CSIs in Longman et al. (2014), Experiment 2. An independent t-test found the RISC slope to be steeper in the present experiment, $t(40) = 3.29$, $p=.002$. However, although numerically smaller by ~30 ms in

the present experiment when compared to the mean of switch costs in CSIs 1020 ms and 1420 ms in Longman et al., (2014), Experiment 2 (a difference that did not approach significance, Switch by Experiment interaction, $F(1,40)=2.04$, $MSE=1808.60$, $p=.16$), the residual RT switch cost was a non-trivial and significant 45 ± 17 ms (38% of the switch cost in the shortest CSI). Thus, we must conclude that whilst self-paced preparation can eliminate "inertia" in the spatial attention parameters it is not sufficient to overcome other source(s) of the residual switch cost.

Why is this? To return to the various candidate sources of the residual cost reviewed in the Introduction, if "inertia" in task-set parameters (Allport et al., 1994; Yeung & Monsell, 2003; Yeung et al., 2006) is responsible for the residual cost in RT, we lack a principled account of why self-paced preparation can completely overcome such inertia for one component – spatial attention – and not others (as reflected in the residual RT switch cost). Note that these "others" include other aspects of attention – there is clear evidence that residual switch costs in RT can arise in part from incomplete reorientation of attention to perceptual dimensions (Elchlepp, Best, Lavric, & Monsell, submitted; Mayr et al, 2013, Meiran & Marciano, 2002; Müller, Reimann, and Krummenacher, 2003), or to lexical versus perceptual attributes (Elchlepp, Lavric & Monsell, 2015), though we do not know whether self-paced stimulus presentation would change this. The same point can be made with respect to the claim that residual costs arise from associative retrieval of competing parameter values (Waszak et al., 2003, 2005): if the presence of the stimulus is required for such associative retrieval, that may explain why the resulting conflict can be dealt with only after the stimulus onset, but that would appear to be true no less for attentional than for other parameters. An analogous argument applies to Mayr et al.'s (2014) similar attribution of switch costs to intrusions from long-term memory representations of competing task-set(s). And if, in De Jong's (2000) terms, the pro-active TSR occasionally "fails to engage" for some

task-set components, but not others, why can this tendency be overcome by self-pacing for spatial attention, but not other task-set components? Finally, although it is straightforward, within a two-stage TSR accounts that distinguishes between pre- and post-stimulus TSR processes (e.g., Meiran et al., 2000; Rogers & Monsell, 1995; Rubinstein et al., 2001), to propose that some parameters (e.g., spatial attention) can be fully re-set before the stimulus onset given optimal preparation conditions, whereas others (e.g., S-R mappings) cannot, this is as arbitrary an assumption as those above. Thus any of the above theories of the residual task-switch cost requires augmentation to account for the present set of results; such augmentation may require data on the impact of self-pacing on the contribution of other aspects of task-set (including attention) to the residual cost.

To conclude, we report here a condition under which it is possible to eliminate the effects of 'attentional inertia' on fixations – giving the participant ultimate control over the duration of the preparation interval. However, the benefits of this opportunity for preparation do not extend to at least some other components of the reconfiguration process (indexed by the performance switch cost and the delay in orienting to the relevant stimulus element).

References

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing*, (pp. 421-452). Cambridge, MA: MIT Press.

Arrington, C. M., Logan, G. D., & Schneider, D. W. (2007). Separating cue encoding from target processing in the explicit task-cuing procedure: Are there "true" task switch effects? *Jopurnal of Experimental Psychology: Learning, Memory and*

Cognition, 33 (3), 484-502.

Astle, D. E., Jackson, G. M., & Swainson, R. (2006). Dissociating neural indices of dynamic cognitive control in advance task-set preparation: An ERP study of task switching. *Brain Research,* 1125, 94-103.

De Jong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds), *Control of Cognitive Processes XVIII: Attention and Performance*, (pp. 357-376). Cambridge, MA: MIT Press.

Dixon, P. (1981). Algorithms and selective attention. *Memory & Cognition*, 9, 177-184.

Dixon, P., & Just, M. (1986). A chronometric analysis of strategy preparation in choice reactions. *Memory & Cognition*, 14, 488-500.

Elchlepp, H., Best, M., Lavric, A., & Monsell, S. (submitted). Shifting attention between visual dimensions as a source of the task switch cost.

Elchlepp, H., Lavric A., Mizon, G., & Monsell, S. (2012). A brain-potential study of preparation for and execution of a task-switch with stimuli that afford only the relevant task. *Human Brain Mapping*, 33, 1137-1154.

Elchlepp. H., Lavric, A., & Monsell, S. (2015). A change of task prolongs early processes: evidence from ERPs in lexical tasks. *Journal of Experimental Psychology: General*, 144 (2), 299-325.

Hubner, R., Futterer, T., & Steinhauser, M. (2001). On attentional control as a source of residual shift cost: evidence from two-component task shifts. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 27, 640-653.

Karayanidis, F., Coltheart, M., Michie, P. T., & Murphy, K. (2003). Electrophysiological correlates of anticipatory and poststimulus components of task switching. *Psychophysiology*, 40, 329–348.

Karayanidis, F., Jamadar, S., Ruge, H., Phillips, N., Heathcote, A., & Frostmann, B. U. (2010). Advance preparation in task switching: Converging evidence from behavioural, brain activation and model-based approaches. *Frontiers in Psychology,* 1, 25.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Phillip, A., & Koch, I. (2010). Control and interference in task switching - A review. *Psychological Bulletin,* 136, 849-874.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science,* 6, 449-467.

Lavric, A., Mizon, A., & Monsell S. (2008). Neurophysiological signature of effective anticipatory task-set control: a task-switching investigation. *European Journal of Neuroscience*, 28, 1016-1029.

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance,* 29, 575-599.

Longman, C. S., Lavric, A., & Monsell, S. (2013). More attention to attention? An eye-tracking investigation of selection of perceptual attributes during a task switch. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 39 (4), 1142-1151.

Longman, C.S., Lavric, A., & Monsell, S. (2016). The coupling of spatial attention and other components of task-set: a task switching investigation. *Quarterly Journal of Experimental Psychology.* Advance online publication. doi: 10.1080/17470218.2015.1115112.

Longman, C. S., Lavric, A., Munteanu, C., & Monsell, S. (2014). Attentional inertia and

    delayed orienting of spatial attention in task switching. *Journal of Experimental*

    Psychology: Human Perception and Performance, 40 (4), 1580-1602.

Mayr, U., Kuhns, D., & Hubbard, J. (2014). Long-term memory and the control of

    attentional control. *Cognitive Psychology*, 72, 1-26.

Mayr, U., Kuhns D., & Rieter, M. (2013). Eye movements reveal dynamics of task

    control. *Journal of Experimental Psychology: General,* 142 (2), 489-509.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance.

    Journal of Experimental Psychology: Learning, Memory and Cognition, 22, 1423-

    1442.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching.

    Cognitive Psychology, 41, 211-253.

Meiran, N., Kessler, Y., & Adi-Japha, E. (2008). Control by action representation and

    input selection (CARIS): A theoretical framework for task switching. *Psychological*

    Research. 72, 473-500.

Meiran, N., & Marciano, H. (2002). Limitations in advance task preparation: Switching

    the relevant stimulus dimension in speeded same-different comparisons. *Memory &*

    Cognition, 30 (4), 540-550.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences,* 7 (3), 134-140.

Monsell, S. (2015). Task-set control and task switching. In J. M. Fawcett, E. F. Risko,

    & A. Kingstone (Eds), *The Handbook of Attention* (pp. 139-172). Cambridge, MA:

    MIT Press.

Monsell, S., & Mizon, G. A. (2006). Can the task-cuing paradigm measure an

    "endogenous" task-set reconfiguration process? J*ournal of Experimental Psychology:*

    Human Perception and Performance, 32, 493-516.

Müller, H. J., Reimann, B., & Krummenacher, J. (2003). Visual search for singleton feature targets across dimensions: Stimulus- and expectancy-driven effects in dimensional weighting. J*ournal of Experimental Psychology: Human Perception and Performance*, 29 (5), 1021-1035.

Nieuwenhuis, S., & Monsell, S. (2002). Residual costs in task switching: Testing the failure-to-engage hypothesis. *Psychonomic Bulletin and Review*, 9, 86–92.

Richter, F. R., & Yeung, N. (2014). Neuroimaging studies of task switching. In J. Grange & H. Houghton (Eds.), *Task Switching and Cognitive Control* (pp. 237-271). New York, NY: Oxford University Press.

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

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 763-797.

Ruge, H., Jamadar, S., Zimmerman, U., & Karayanidis, F. (2013). The many faces of preparatory control in task switching: reviewing a decade of fMRI research. *Human Brain Mapping*, 34 (1), 12-35.

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, 136, 601-626.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology*, 46, 361-413.

Waszak, F., Hommel, B., & Allport, A. (2005). Interaction of task readiness and automatic retrieval in task switching: Negative priming and competitor priming. *Memory and Cognition*, 33 (4), 595-610.

Wylie, G. R., Javitt, D. C., & Foxe, J. J. (2006). Jumping the gun: Is effective
preparation contingent upon anticipatory activation in task-relevant neural circuitry?
Cerebral Cortex, 16, 394-404.

Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: the
role of stimulus-attribute and response-set selection. *Journal of Experimental
Psychology: Human Perception and Performance, 29, 455-469.*

Yeung, N., Nystrom, L. E., Aronson, J. A., & Cohen, J. D. (2006). Between task
competition and cognitive control in task switching. *Journal of Neuroscience,* 26 (5),
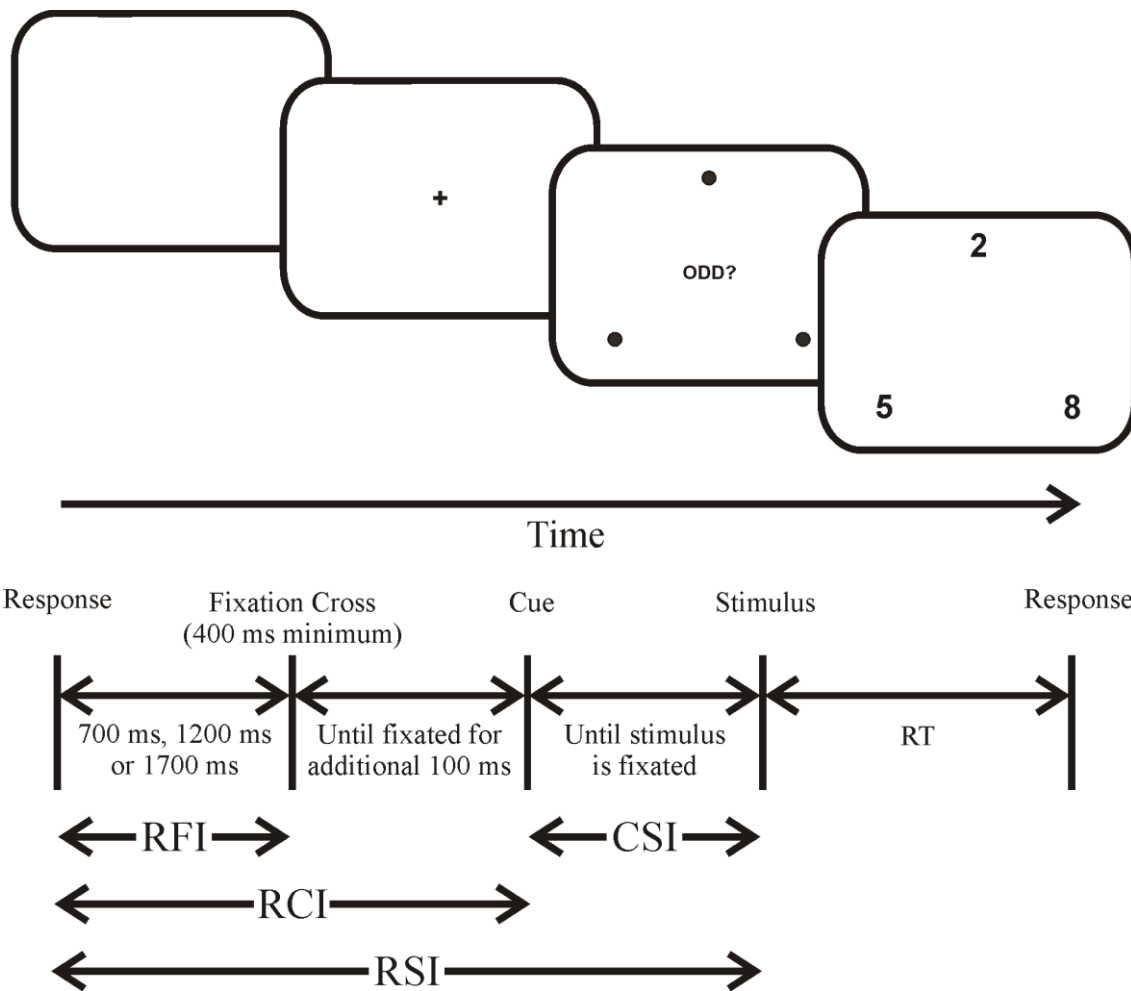1429-1438.

Figure 1. The time-course of one trial with example displays. The dots used to identify where the digits would eventually appear were presented in light blue. Note that each frame shows only the central area of the screen. RFI = response-to-fixation cross interval, RCI = response-to-cue interval, RSI = response-to-stimulus interval, CSI = cue-to-stimulus interval.
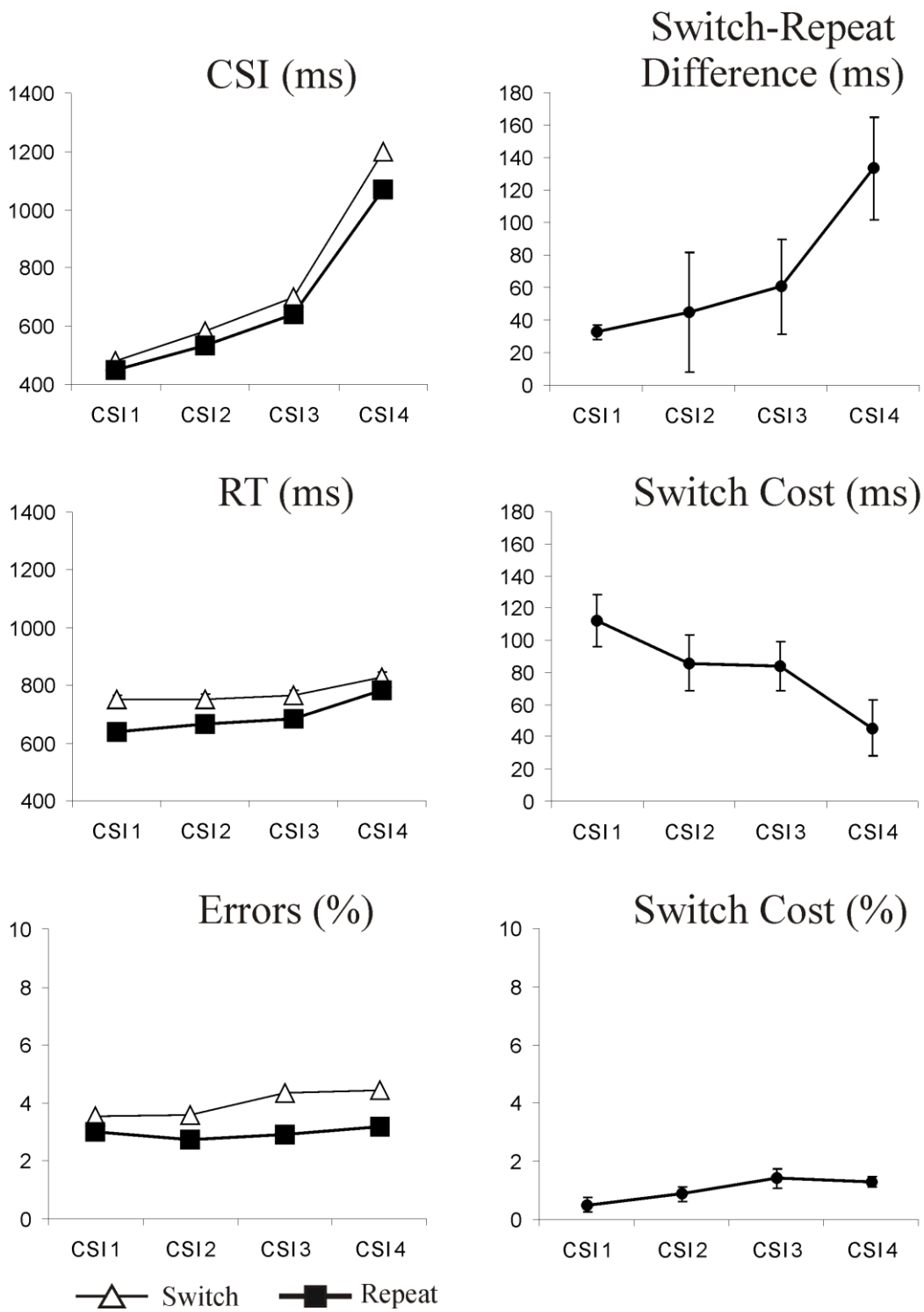
Figure 2. Mean cue-stimulus interval (CSI; top), RT (middle) and error rate (bottom) as a function of CSI quartile and transition (switch, repeat). The column on the right shows the (switch-repeat) difference for the CSI durations and switch costs for the behavioral data (error bars show the standard error of the mean difference).
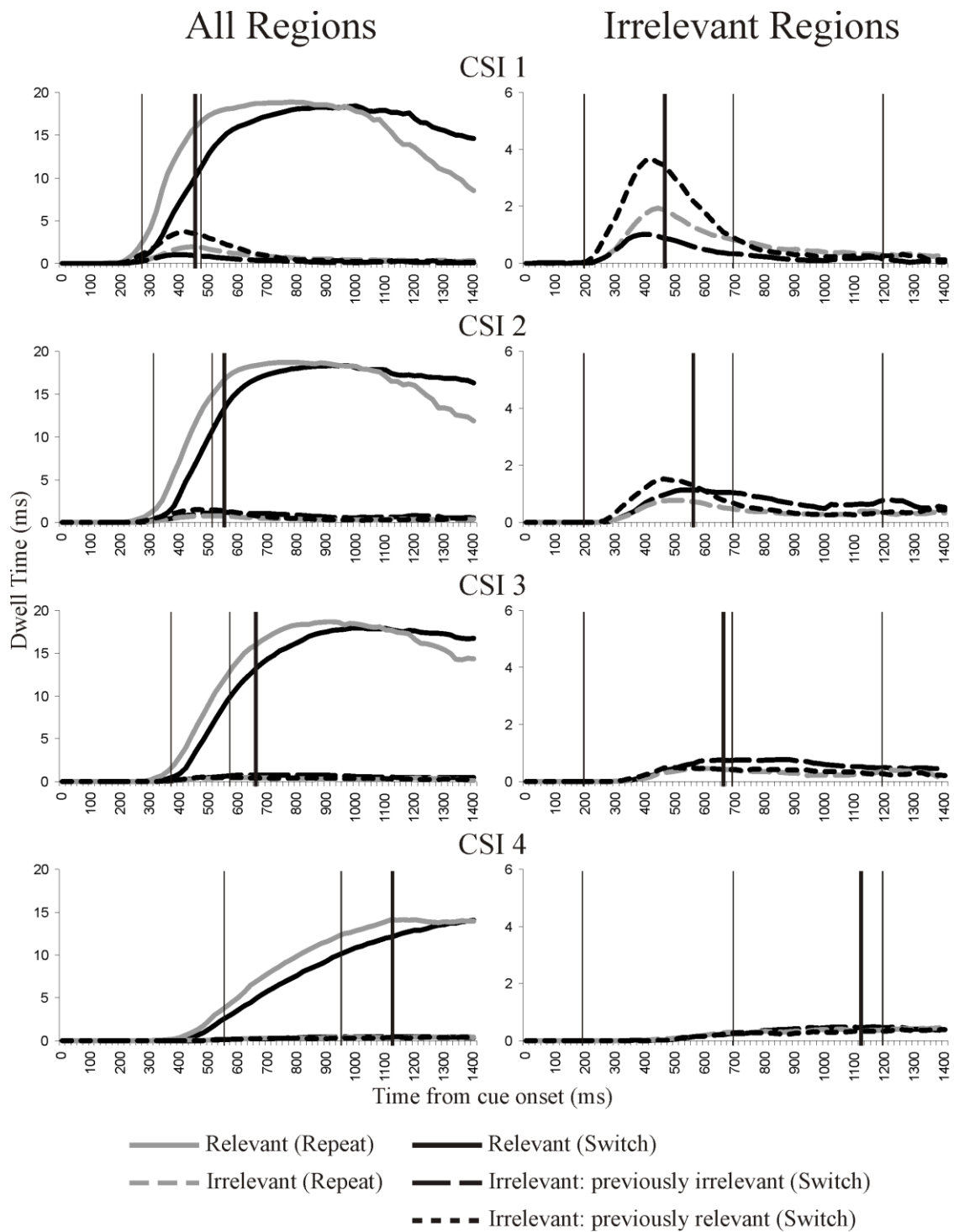
Figure 3. Dwell time per 20 ms (see text for definition) for 1420 ms following cue onset as a function of CSI quartile, transition (switch, repeat) and task-relevance of stimulus region. The left panel shows the fixations on both the relevant and irrelevant regions, whereas the right panel shows only the fixations on the irrelevant regions (with the scale adjusted accordingly). Thick vertical lines indicate the mean stimulus onset time, thin vertical lines show the time-windows used for statistical analysis.

Table 1.

Mean (Standard Error) Difference (Previously Relevant-Previously Irrelevant) in Dwell Time in the Currently Irrelevant Regions on Switch Trials in the Early and Late Time Windows as a Function of CSI and RFI.

	RFI	CSI 1		CSI 2		CSI 3		CSI 4	
Early Window	**700**	28.74	(9.95)	5.21	(7.5)	-0.69	(1.81)	0.18	(0.77)
	1200	22.43	(9.02)	5.53	(5.82)	-3.49	(1.4)	-0.3	(1.05)
	1700	23.44	(9.13)	4.87	(4.87)	-0.37	(1.64)	0.07	(0.6)
	Mean	24.87	(8.83)	5.2	(5.57)	-1.51	(1.17)	-0.02	(0.58)
Late Window	**700**	-3.75	(2.52)	-10.07	(4.93)	-6.15	(3.25)	-1.28	(2.54)
	1200	-1.66	(3.08)	-3.69	(5.18)	-13.21	(2.82)	-3.37	(2.76)
	1700	-3.65	(3.04)	-12.41	(4.84)	-1.97	(4.57)	-4.13	(1.94)
	Mean	-3.02	(1.81)	-8.72	(3.47)	-7.11	(2.84)	-2.93	(1.48)