Eye Movements Reveal Dynamics of Task Control

Ulrich Mayr and David Kuhns
University of Oregon

Miranda Rieter
University of Colorado at Boulder

With the goal to determine the cognitive architecture that underlies flexible changes of control settings, we assessed within-trial and across-trial dynamics of attentional selection by tracking of eye movements in the context of a cued task-switching paradigm. Within-trial dynamics revealed a switch-induced, discrete delay in onset of task-congruent fixations, a result that is consistent with a higher level configuration process. Next, we derived predictions about the trial-to-trial dynamic coupling of control settings from competing models, assuming that control is achieved either through task-level competition or through higher level configuration processes. Empirical coupling dynamics between trial $n-1$ eye movements and trial n response times—estimated through mixed linear modeling—revealed a pattern that was consistent with the higher level configuration model. The results indicate that a combination of eye movement data and mixed modeling methods can yield new constraints on models of flexible control. This general approach can be useful in any domain in which theoretical progress depends on high-resolution information about dynamic relationships within individuals.

Keywords: executive control, task switching, eye movements, mixed modeling

Is flexible, goal-directed action implemented in a hierarchical manner and through higher level configuration processes? Or is it achieved through local competition between representations related and unrelated to the current goal? This long-standing issue in cognitive psychology (e.g., Botvinick & Plaut, 2004; Cooper & Shallice, 2006) is relevant for questions such as how to characterize developmental (Bunge & Zelazo, 2006) and disease-induced changes in cognitive control (Waltz et al., 2004), or for how to establish correspondence between behavioral and neural-level data on the neurocognitive architecture of cognition (Sporns, Chialvo, Kaiser, & Hilgetag, 2004). In the present work, we approach this topic by drawing on three different literatures: (a) research on task switching in which the question about the nature of the control architecture has recently come to the forefront (e.g., Gilbert & Shallice, 2002; Schneider & Logan, 2005); (b) Research that has established eye tracking as a tool for characterizing allocation of attention in a fine-grained manner (Theeuwes, 2010); (c), recent developments in statistics—mixed, linear modeling—that allow characterizing the dynamics of control within individual subjects (e.g., Bates, Maechler, & Dai, 2008; Gelman & Hill, 2007; Kliegl, Masson, & Richter, 2010).

In the next section, we present how the task-switching paradigm and eye tracking can be combined to assess control dynamics in a fine-grained manner (see the Task Switching and Eye Movements section). Then, we introduce the hierarchical configuration model and two different variants of "nonhierarchical" models (see the Models of Task Selection section). We explain how these models can be tested by looking at the dynamics of task selection within trials (see the Within-Trial Control Dynamics section) and by examining the dynamic coupling of control settings across trials, using mixed, linear modeling techniques (see the Between-Trial Dynamic Coupling of Control Settings section). In the first part of the Results section, we demonstrate how within-trial dynamics of eye movements provide initial evidence in support of the configuration model (see the Results: Within-Trial Dynamics section). In the second part of the Results section, we show that the analysis of trial-to-trial dynamics provides additional, independent information that largely supports the configuration model (see the Results: Trial-to-Trial Coupling Dynamics section).

Task Switching and Eye Movements

The task-switching paradigm has become an important tool for tackling the question of how people achieve flexible control over cognitive operations (Logan, 2003; Meiran, 1996; Monsell, 2003). Subjects work with two or more response time tasks (e.g., judging the color or the shape of an object). Either an external cue or a sequential rule specifies which task is relevant on a given trial. The response time (RT) difference between trials after a task switch and after a task repeat, the switch cost, is used to capture the demands of changing task sets between successive trials. Switch costs can be substantial (i.e., 150–300 ms) when there is little time to prepare and are reduced when time for preparation is provided. However, proactive control rarely eliminates switch costs completely, even when people have ample opportunity to prepare.

There are important theoretical differences between competing models of task switching, some of which we turn to below. However, all accounts agree that task selection costs directly reflect interference that arises from currently irrelevant tasks, or

the processing operations that are necessary to fight this interference (or both; see Kiesel et al., 2010; Vandierendonck, Liefooghe, & Verbruggen, 2010). Therefore, a precise characterization of the temporal interference/control dynamics is key to the understanding of flexible task control.

With standard RT measures of task switch performance, we only obtain a static assessment of interference based on difference scores between condition averages (e.g., no-switch vs. switch RTs). In contrast, by tracking eye movements during task switching, we can trace for a given individual, trial, and even time point within a trial, whether attention is either on task-relevant or on task-irrelevant aspects. This should get us much closer to a detailed assessment of real-time control and interference dynamics.

In order to use eye tracking in a task selection context, we need to separate task-relevant features across different objects on the screen rather than combining them within the same object, as it is typical in task-switching situations. Figure 1 demonstrates key features of our paradigm. We used two possible tasks, the "color" and the "gap" task, which were indicated through verbal task cues. Cues were presented either 300 ms or 1,000 ms prior to the stimulus to manipulate opportunity for proactive control. On each trial, three vertical bars were displayed, located equidistantly from each other on a virtual circle. One of the three bars differed in color from the rest; it was dark or light blue, whereas the others were medium blue. A second bar had a gap in it, either at the top or at the bottom of the bar, whereas the others had no gap. Subjects were asked to respond either to the color (light blue = left response, dark blue = right response) or to the gap (bottom gap = left response, top gap = right response). We refer to the task-relevant bar as the *target* and to the other singleton as the *distractor* because it is task relevant on other trials. The third bar is never

relevant, and therefore we refer to it as the *neutral* object. Except for their distance from fixation and the configuration in an equilateral triangle, object locations were unpredictable. Thus, potential trial-to-trial effects on eye movements could not be interpreted in terms of location priming.

To perform either task, [*]subjects had to (a) attend to the appropriate object and (b) apply the relevant S-R rules to that object. We can capture the first process, task-congruent attention, by looking at the probabilities of target- and distractor-directed fixations. Even though there is no complete overlap between eye movements and attention, fixation information can serve as a reliable indicator of where attention is directed to in most unconstrained situations (e.g., Awh, Armstrong, & Moore, 2006). Research on eye movements and attention has established important benchmark results of how fixations respond to bottom-up signals, to top-down control attempts, and to global strategic settings (e.g., Dodd, Van der Stigchel, & Hollingworth, 2009; Findlay, 1997; Van der Stigchel, Meeter, & Theeuwes, 2006; van Zoest, Donk, & Theeuwes, 2004; for a different view, see Folk & Remington, 2006). In particular, within the first 200 + ms after stimulus display, fixations are attracted to low-level visual features (e.g., singleton objects) in an unconditional manner, and only thereafter do top-down control settings begin to take hold (van Zoest et al., 2004; see also Theeuwes, 2010). Thus, target versus distractor probabilities as a function of time should allow us to determine a particularly important time point, namely the onset of task-congruent, top-down control. As discussed in the next sections, models of task selection differ in their predictions about how this onset of control is affected by switches in task sets.

Although eye movements inform about attentional selection, they do not provide *direct* information about the second process,
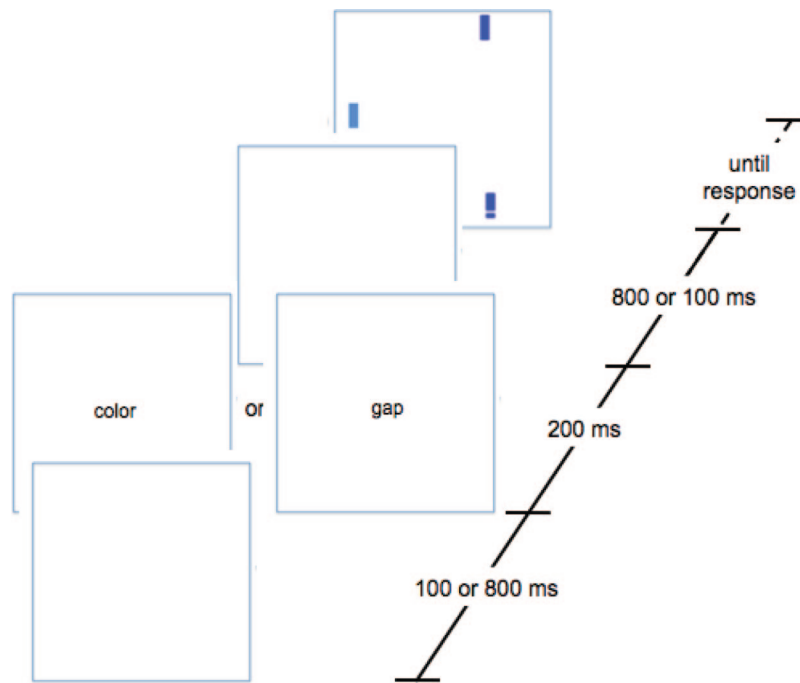

Figure 1.   Successive stimulus displays for a single trial. Note that two possible cues could be used per task. Stimuli are not drawn to scale.

the selection of relevant S-R rules. This is an important issue that we return to in the General Discussion (see the Qualifications section).

Models of Task Selection

Most models of task control are situated within the theoretical tension between two polar opposites, which we refer to here as *configuration models*, on the one hand, and *task-level selection models*, on the other. Configuration models assume that a switch in task necessitates a process (or a set of processes) that operates on an abstract, hierarchically higher level than that of specific tasks. Such processes may involve the retrieval of task rules from long-term memory into working memory (e.g., Logan & Gordon, 2001; Mayr & Kliegl, 2000, 2003; Monsell & Mizon, 2006) or inhibition of the no-longer relevant task (e.g., Mayr & Keele, 2000).

The following additional assumptions will be critical here: (a) At least in part, configuration requires a time-limited control operation that needs to be completed before task-specific processing can begin (e.g., Meiran, 1996; Rogers & Monsell, 1995). (b) In principle, configuration can be executed proactively (i.e., prior to stimulus presentation; Meiran, 1996). (c) However, previous work also suggests that whether or not, or when exactly (de Jong, 2000), configuration takes place depends on strategic settings subjects adopt in response to the global or local selection context. For example, when overall switch frequency is high, people often tend to lower their threshold for initiating configuration processes, such as inhibition of the previous task and updating of working memory, so that these can occur even on no-switch trials (e.g., Mayr, 2006; Monsell & Mizon, 2006). Thus, although in principle configuration is mandatory on switch trials, it can be—but is not always—avoided on no-switch trials. (d) A final assumption that will become critical when we turn to the between-trial dynamic is that the strength of high-level control is not constant, but rather fluctuates over time (e.g., West & Travers, 2008).

Does flexible switching between competing tasks actually require higher level control and reconfiguration? Kiesel et al. (2010) recently concluded a review of relevant research by stating that it is not clear whether it is ". . . theoretically necessary to postulate executive control processes to explain switch costs" (p. 868). Specifically, configuration models need to earn their credibility against a category of models that assume a single process for no-switch and switch transitions and/or a "flat control hierarchy," where control is established strictly through task-specific competition and where trial-to-trial carryover of task-specific representations is responsible for switch costs. There are two different models, the compound-cue model by Logan and colleagues (e.g., Logan & Bundesen, 2003; Schneider & Logan, 2005) and the connectionist task-level selection model (Gilbert & Shallice, 2002), that have been designed explicitly to function without higher level, switch-related processes. We focus on these two models because, albeit different in their explanations, they represent the strongest theoretical case in favor of "nonhierarchical" accounts of task switching.

We do not want to ignore here that one way to solve the theoretical tension between the opposing models is through hybrid models that include both high-level and low-level selection processes (e.g., Brown, Reynolds, & Braver, 2007; O'Reilley, 2006). However, presently there are no good diagnostics available to clearly identify either configuration or carryover dynamics in empirical data. Therefore, predictions from pure models are a good starting point for developing such diagnostics.

Within-Trial Control Dynamics

One straightforward prediction of a model that assumes an initial configuration operation is that such an operation would postpone the onset of task-congruent attention in a relatively discrete manner on switch trials relative to no-switch trials. Also, people should be able to initiate configuration proactively, which implies that the switch-related discrete delay in task-congruent eye movements ought to disappear when opportunity for preparation is provided. Furthermore, the delay in onset of task-congruent attention on switch trials should be dependent on context. Specifically, if the probability of switches is high, people should be more inclined to initiate reconfiguration or inhibition even on no-switch trials (e.g., Mayr, 2006; Monsell & Mizon, 2006). As a result, the discrete delay in the onset of control should occur for switch and no-switch trials alike.

Alternative predictions about within-trial control dynamics can be derived from the above-mentioned compound-cue model (Logan & Bundesen, 2003; Schneider & Logan, 2005). According to this model, subjects in cued task-switching situations do not operate with *task-specific* attentional settings, but rather with one unified task set that integrates all relevant rules. Switch costs arise because only on no-switch trials are task cue representations carried over from the previous-trial cue. In the standard task-switching paradigm, cues on successive trials are identical in case of a no-switch transition, thus producing strong cue-priming effects. However, even when two cues per task are used to avoid such cue repetitions (e.g., Mayr & Kliegl, 2003), the two cues associated with the same task are semantically or, at least episodically, related, thus potentially allowing for some priming. Once the task cue is encoded, it serves, combined with the target, as a "compound cue" that disambiguates the correct response. The compound-cue model further assumes that switch costs arise from exponential distributions of the time it takes to finish cue-encoding processes that differ for trials on which compound cues are similar to those on the preceding trial (i.e., no-switch trials) relative to trials on which compound cues change (switch trials). Because these functions differ only in the rate at which response-relevant information accumulates (depending on the amount of priming), the model predicts a *common onset* for task-congruent attentional processing on switch and no-switch trials. In other words, the compound-cue model has no mechanism for explaining a potential discrete delay in the onset of task-congruent attention on switch trials. Moreover, given the assumption that task-specific attentional settings play no role in cued task-switching situations, any finding of switch effects on task-congruent attention would be problematic for this model.

In the first part of the Results section, we report data on within-trial task selection dynamics. Some readers may be mainly interested in these aspects of the results, which can be adequately evaluated even when skipping the following section on between-trial dynamics.

Between-Trial Dynamic Coupling of Control Settings

As mentioned earlier, fixation information provides an index of task-relevant selection that can be used in a trial-by-trial manner. This in turn allows us to characterize trial-by-trial control dynamics as a new source of information to evaluate models of control. Specifically, we look at how attentional processing on trial $n - 1$ predicts attentional processing and/or RTs on trial n. Our goal in the next sections is to derive from each of the two competing types of models specific predictions about trial-to-trial control dynamics, to verify these predictions via simple model simulations, and to establish how best to test these in both simulated and actual data.

Control Dynamics Generated Through Task-Level Selection Models

Figure 2a is an abstract depiction of the model of task selection by Gilbert and Shallice (2002) in which attentional control oper-

Figure 2. Schematic representation of (a) the task-level selection model (Gilbert & Shallice, 2002) and (b) the higher level configuration model. Arrows indicate positive relationships; oval line endings indicate negative (i.e., inhibitory) relationships. Arrows that return to a unit represent carryover of activity across trials. Red and green arrows denote relationships that need to change on switch trials.

ates on the level of tasks. On each trial, selection is achieved through local competition between the cued task and the currently nonrelevant task. This competition takes place through lateral inhibition between so-called task demand units, which biases attentional processing and/or response selection. On switch trials, selection takes longer because the activation level in the task demand units carries over from one trial to the next (i.e., symbolized through the paths from each task demand unit to itself). This gives the formerly appropriate, but now inappropriate demand unit a head start that needs to be overcome before processing of task-relevant information can begin.

Note, that in this model the only change from trial to trial is the "task input" (e.g., in form of task cues) that indicates which task is currently relevant. Otherwise, the network does not have to be "reconfigured" on switch trials. Thus, task selection effects arise exclusively from the basic "associative machinery," plus the assumption that task demand activity carries over across trials. This model implements the intuitively appealing idea that inertia of task-specific attentional settings from one trial to the next can produce important mean-level task-switching phenomena without requiring assumptions about higher level, task-unspecific control processes (for details, see Gilbert & Shallice 2002, and the Appendix).[1] However, from this model we can also derive specific predictions about the trial-to-trial dynamic coupling between successive control states. To make these dynamics transparent, we have rerepresented the model in Figure 3. The way to read these panels is to think of them as path-analytic regression models, where trial $n-1$ variables are used to predict trial n dependent variables. For now, we focus on the top-left panel of this figure; the remaining panels are explained below. To arrive at a succinct representation of the control dynamics, we need to abstract away from the specific tasks (i.e., color or gap), and, rather, distinguish between the relevant and the irrelevant tasks for each trial (i.e., corresponding to the "task-relevant control" and "task-irrelevant control" units). Although with simulations (see the Appendix) we can directly probe activity in the task demand units, in behavioral data the status of task-specific control settings can be assessed only via proxies. In Figure 3, the labels "target" and "distractor" represent these proxies. In our case, they stand for binary predictors indicating whether or not on that trial an eye movement was made to the target or distractor. Accordingly, the figure distinguishes between hypothetical (i.e., dotted lines) and empirically observable relations (i.e., filled lines). In theory, of course, the observable
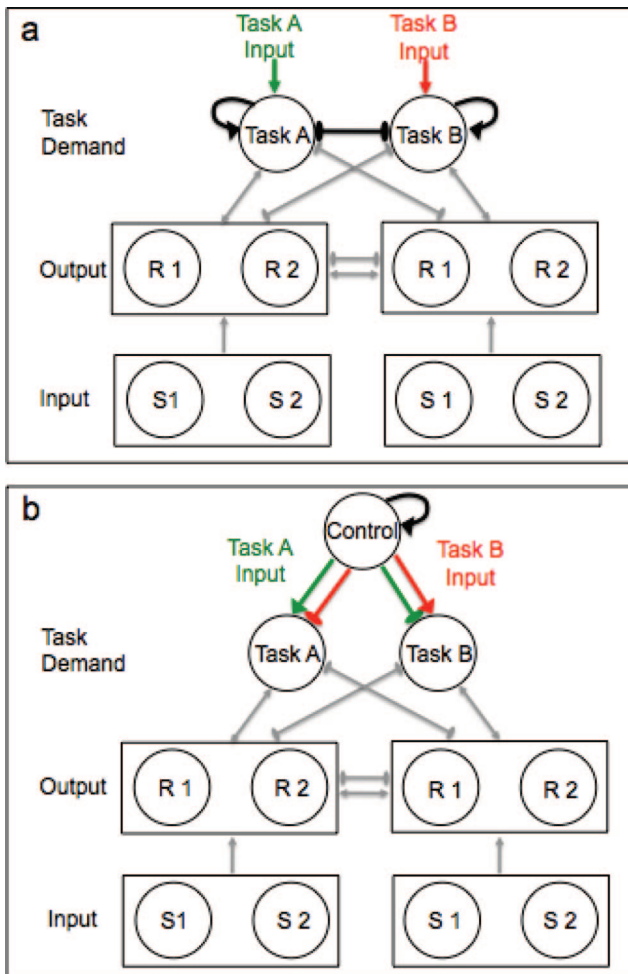
---

[1] One might argue that within the Gilbert and Shallice (2002) model, the "task input" that specifies the currently relevant task can be conceived as a form of "higher level control." On the one hand, we agree that in principle, the task inputs could serve this function. For example, if one wanted to augment this model through a task-unspecific level of control (i.e., without changing its current task-specific selection mechanism), the best way to do this would be by explicitly specifying how the task input is computed. On the other hand, in the original model the task input contributes nothing to the explanation of task switch costs or trial-to-trial dynamics (e.g., it is identical for switch vs. no-switch trials). Switch costs arise from the task selection dynamics that take place within the "associative machinery" once all stimulus and task input parameters are provided. Thus, although one could in principle conceive this model as containing a "higher," task-unspecific level of selection, this level has no functional relevance for what we try to explain here.
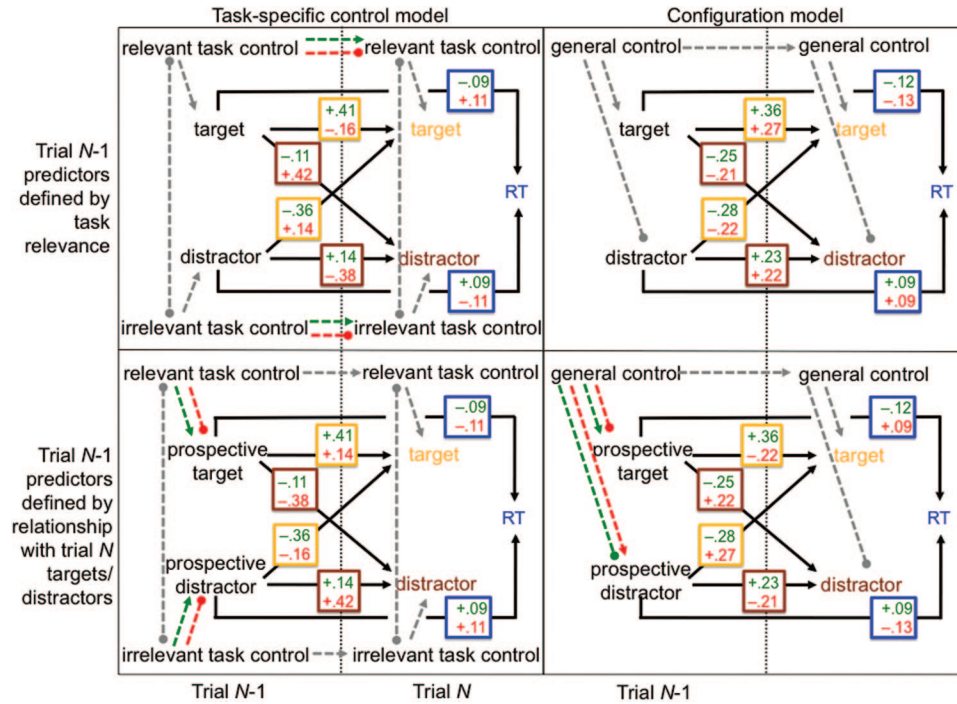
Figure 3. Representation of the task-level selection model (left two panels) and the higher level configuration model (right two panels), along with two complementary analyses of trial-to-trial control dynamics: categorizing trial $n-1$ eye movements in terms of relevance (top panels) and categorizing trial n-movements relative to the "prospective" relevance on trial n (bottom panels). Dotted lines represent hypothetical and filled lines represent observable relationships. Lines, symbols, or coefficients in green are relevant for no-switch transitions, those in red for switch transitions, and gray lines are assumed to be invariant across no-switch and switch transitions. *Target* and *distractor* refer to indicators of attention (e.g., eye movements) to the target or distractor object on that trial (or to prospective targets/distractors in bottom panels). To avoid overcrowding of the figure, we omitted inhibitory links between the relevant task control and distractors and between irrelevant task control and targets in the top and bottom left panels. Coefficients and dependent variables associated with the same model are contained in boxes with the same color (Model 1: orange; Model 2: brown; Model 3: blue). Numbers in boxes are correlation coefficients and show the results from a simulation of the Gilbert and Shallice (2002) task-level selection model (left panels) and the configuration model (right panel). RT = response time.

relations are constrained by the status of the hypothetical variables and the relations between them.

So how should the trial-to-trial coupling dynamics play out? For no-switch trials, the task control unit relevant on trial $n-1$ is the same as on trial n. Thus, whatever the activity level on trial $n-1$ might be, it will carry over into trial n (expressed through the excitatory links between corresponding task demand units across trials in top-left panel of Figure 3). This in turn implies a positive correlation for the "horizontal relationships" between trial $n-1$ and trial n eye movements (i.e., $n-1$ relevant $->$ n relevant and $n-1$ irrelevant $->$ n irrelevant). Furthermore, the mutual inhibition on the level of task demand units will enforce that the "diagonal relations" (i.e., $n-1$ irrelevant $->$ n relevant, and $n-1$ irrelevant $->$ n relevant) are negative. In other words, for no-switch trials, the efficiency of the control setting on trial $n-1$ will be generally positively related to the efficiency of the control setting on trial n. In contrast, on switch trials, the specific task that was relevant on trial $n-1$ becomes irrelevant on trial n, and vice versa. In Figure 3, this is expressed through the horizontal inhibitory links between trial $n-1$ and trial n task control units for

switch trials. On the level of observable relationships, this implies the opposite pattern than for no-switch trials, namely negative "horizontal" and positive "diagonal" relations between trial $n-1$ and trial n eye movement indicators of attentional efficiency.

To confirm these predictions, we adapted a version of the Gilbert and Shallice (2002) model to our experimental situation (see correlation coefficients in top-left panel of Figure 3 and the Appendix for modeling details). Modeling results show that the horizontal relationships (i.e., $n-1$ target $->$ n target, $n-1$ distractor $->$ n distractor) are positive on no-switch trials, whereas the diagonal relationships are negative (i.e., $n-1$ target $->$ n distractor or $n-1$ distractor $->$ n target). However, on switch trials, we see exactly the opposite pattern. For example, the correlation between the trial $n-1$ target and the trial n target flips from .41 for no-switch trials to $-.16$ for switch trials. Equally, attention to the trial $n-1$ target predicts faster trial n RTs (i.e., $-.09$) on no-switch trials, but slower RTs on switch trials (i.e., $+.11$), whereas the reverse relationship holds for the relationship between trial $n-1$ distractors and trial n RTs (i.e., .09 vs. $-.11$). Combined, the simulation results illustrate the basic prediction of the carryover

model for trial-to-trial control dynamics: The more efficient the trial $n-1$ attentional setting, the more efficient the setting is on trial n in case of a no-switch transition, but the less efficient it is following a switch transition.

Control Dynamics Generated Through Higher Level Configuration

In Figure 2b, we changed the task-level model into a hierarchical control model. To this end, we replaced the direct competition between task demand units by a task-unspecific, general control unit that can selectively activate or deactivate the task-specific demand units, depending on which task is currently relevant. Thus, whereas in the task-specific model the "associative wiring pattern" remains constant across no-switch and switch trials, in the configuration model the wiring pattern itself needs to be changed on switch trials (represented by the green and red arrows in Figure 2b). To account for the waxing and waning of control over time (e.g., West & Travers, 2008), activity in the higher level unit carries over between trials (again symbolized by the paths from the high-level control unit to itself in Figure 2b; see the Appendix). Thus, whereas in the task-level selection model control settings carry over in a task-specific manner, in the hierarchical control model, carryover occurs on the level of a higher level control resource that fuels the task-specific selection processes.

The purpose of the model shown in Figure 2b is to propose an operational definition of the minimal difference between the task-unspecific and the task-specific control model and to show how from this difference specific dynamic coupling patterns may arise. However, we acknowledge that this abstract model representation ignores the hard problem of specifying the mechanism for flexibly "rewiring" the links between the task-unspecific control unit and the task demand units, a problem that does not exist in the Gilbert and Shallice (2002) model. Thus, going from the task-specific to the task-unspecific control model clearly comes with a loss in parsimony. The more important it is to examine whether with this change come unique predictions and greater explanatory power.

In fact, the configuration model makes different predictions about the trial-to-trial coupling pattern than the task-specific control model. The reason is that the trial-to-trial variations of control efficiency are generated "above" the level of specific tasks and therefore do not depend on whether or not tasks change across trials. Thus, the efficiency of control on trial $n-1$ should be positively coupled with efficiency of control on trial n, no matter whether there is a switch between successive trials. In the top-right panel of Figure 3, this is demonstrated by the fact that the horizontal links between trial $n-1$ and trial n relevant and irrelevant task control units are excitatory for no-switch and switch trials alike. Accordingly, on the level of indicators of attentional settings (i.e., eye movements to targets or distractors), the model predicts *identical* trial-to-trial relations for no-switch and switch trials. In fact, simulations with our implementation of the general control model confirm this prediction (see the Appendix for details). Obviously, this pattern contrasts with the switch-dependent flip in coefficients that is produced by the task-specific control model (see top-left panel in Figure 3).

A problem with testing the configuration model in this manner is that it predicts no differences in relationships for no-switch versus switch trials and is thus based on confirmation of the null

hypothesis. Therefore, we also use the alternative specification, shown in the bottom-right panel of Figure 3, which allows a falsification test of the configuration model. The specifications differs in how exactly trial $n-1$ eye movements are used to predict eye movements and RTs on trial n. Whereas the specification in both top panels of Figure 3 categorizes eye movements in terms of task relevance for trial $n-1$ (i.e., target vs. distractor on trial $n-1$), this new specification categorizes trial $n-1$ eye movements in terms of their relationship to what is target and what is distractor on trial n. To be specific: Assume gap is the target and color the distractor on trial n. In this case, we refer to gap as the "prospective target" and color as the "prospective distractor" on trial $n-1$. Let us now further assume attention is directed to the gap dimension on trial $n-1$, which would be an indication of an efficient, goal-directed control setting. If trial n is a no-switch trial, the carryover of general control activity ensures perseveration of the tendency to attend to the gap dimension into the next trial, just as the carryover of the control activity does in the task-level control model (in fact, both specifications make qualitatively identical predictions for no-switch trials). However, if there is a switch between trial $n-1$ and trial n, the carryover of an efficient, general control setting from trial $n-1$ biases the system toward a similarly efficient, but now reconfigured setting that favors attending the (previously ignored) color and ignoring the (previously attended) gap dimension. In the figure, the reconfiguration between trials that needs to happen on switch trials is indicated through the different links (red vs. green) for no-switch and switch trials between the general control unit and trial $n-1$ prospective target/distractor. Due to this reconfiguration, the positive relationship between the prospective target and the trial n target is eliminated or even turns negative for switch trials; the same is true for the relationship between the prospective distractor and the trial n distractor. Thus, with this specification, we explicitly test the prediction that an attentional setting is not carried over for switch trials. Instead it is counteracted by a configuration operation that is fueled by general, task-unspecific control activity and that waxes and wanes across trials. It would be particularly interesting if we find switch-trial coefficients that are not only reduced, but opposite in sign to those on no-switch trials. Such a pattern would indicate that inhibition is used during a task switch to counter previous-trial influences.

In the bottom-right panel of Figure 3, we included results from a simulation using the configuration model (see Figure 2b and the Appendix) with the prospective target/distractor specification. The coefficients show both the predicted carryover pattern for no-switch trials and its reversal for switch trials. For example, for RTs as the dependent variable, trial n RT decreased as attention to the prospective target increased (i.e., $-.12$), whereas for switch trials the reverse relationship emerged: Eye movements to the prospective target lead to slower RTs on the following trial (i.e., $+.09$).

A noteworthy feature of the way we specified the configuration model is that although the currently relevant task receives positive activation, the currently irrelevant task set is actively suppressed. To examine how between-trial coupling dynamics are affected by such active suppression, we tested an activation-only variant of the model shown in the bottom-right panel in Figure 3. The currently irrelevant task was simply allowed to decay to baseline between trials. In this case, the influences from the trial $n-1$ prospective target were qualitatively similar to those shown in Figure 3 (bottom-right panel) on no-switch trials. In contrast, on switch

trials rather than flipping in sign, the corresponding relationships reverted to zero. The opposite pattern emerged for prospective distractors. On no-switch trials, they had no influence on trial n processing. However, on switch trials, where the prospective distractor corresponds to the trial $n-1$ target, the relationships were qualitatively similar to those shown in the bottom-right panel of Figure 3. Thus, an empirical pattern that contains an actual flip in signs (i.e., with significant coefficients in either direction) would be consistent with a configuration process that uses active inhibition of the irrelevant task.

For sake of completeness, the specification in terms of trial $n-1$ prospective targets/distractors can also be applied to the task-level selection model (see bottom-left panel of Figure 3). The task-level model predicts here that whatever dimension is attended on trial $n-1$ will also be attended with greater than chance likelihood on trial n, no matter whether there is a switch between trials or not. Accordingly, the absence of switch effects on coefficients (see Figure 3, bottom-left panel) can be seen as a confirmation of this model, albeit a relatively weak confirmation that, analogously to the situation in the top-right panel of Figure 3, relies on accepting the null hypothesis.[2] Thus, although the appropriate test of the task-level selection model is through the specification in terms of task-relevant and task-irrelevant predictors (top-left panel), the appropriate test of the general control model is through the specification in terms of prospective targets/distractors (bottom-right panel).

In Experiment 2, we tested the competing predictions of the task-level selection and the configuration model. It is important to keep in mind that the trial-to-trial dynamics relevant for these tests occur within individuals. Via mixed linear modeling, we can assess within-individual predictive relationships from trial $n-1$ to trial n, while also accounting for between-individual differences in such relationships (e.g., Bates et al., 2008; Gelman & Hill, 2007; Kliegl et al., 2010). Therefore, the availability of these techniques—at least within experimental psychology, a relatively recent development—is an essential prerequisite for our model testing approach.

Experiments 1 and 2

In Experiments 1 and 2, we used the paradigm shown in Figure 1, which are presented together. Each of the two tasks was cued with two possible labels, such that in alternating trials, only one set (e.g., "color" and "gap") or the other set ("hue" and "space") was presented. This procedure avoids cue repetitions and therefore reduces the effects of cue priming (Logan & Bundesen, 2003; Mayr & Kliegl, 2003). Following Monsell and Mizon (2006), in Experiment 1 we used a 25% switch rate. Arguably, low switch rates induce subjects to maintain tasks across no-switch trials so that these can serve as an appropriate baseline for observing both switch costs and preparation effects.

In addition to a 25% switch rate, Experiment 2 also included switch rates of 50% and 75%, manipulated between subjects. Again, this is based on Monsell and Mizon (2006), who had shown that both switch costs and the preparation effect on switch costs generally decreased with increasing switch frequency. From the perspective of configuration models, this pattern can be explained by assuming that people flexibly adapt their use of configuration processes according to switch frequency. In particular, as switch

frequency increases, subjects' threshold for engaging in reconfiguration processes should decrease. As a result, they should be more likely to initiate reconfiguration in an indiscriminant manner, even on no-switch trials. This makes the strong prediction that for high-switch frequencies, a discrete delay in onset of task-congruent attention is present for no-switch and switch trials alike. In contrast, the Gilbert and Shallice (2002) model has no process for explaining such a pattern.

There were a few additional differences between Experiments 1 and 2: In Experiment 1, we used eye movement responses to measure RTs. This has the disadvantage that at some point during the production of a response, control of fixation, initially directed to stimulus acquisition, is grabbed by the response system, which makes unambiguous interpretation of eye movements as indicators of attention difficult. So, in Experiment 2 we used standard keypress responses. In Experiment 2, but not in Experiment 1, we used incentives to motivate subjects. Finally, only in Experiment 2 did we use a sample size that was sufficient to use mixed modeling for the analysis of across-trial relationships between eye movements or eye movements and RTs.

Method

Subjects. Thirteen students from the University of Oregon participated in Experiment 1; 72 participated in Experiment 2 in exchange for course credits.

Tasks, stimuli, and procedure. Stimuli were presented on a 17-in. Dell CRT monitor set to 1024×768 resolution. Eye movements were recorded using the SR Research desk-mounted Eyelink 1000, controlled by the Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002).

On each trial, three vertically oriented bars were presented, each $1.7°$ high and $.34°$ wide. One bar differed in color from the other two bars (light blue or dark blue instead of medium blue), a second bar had a small gap near either the top or the bottom. Depending

---

[2] Although the carryover model does actually produce no switch effects with the specification in terms of prospective targets/distractors for trial n RTs as dependent variable, matters are a bit more complex for trial n targets or distractors as dependent variables. For example, for the relationship between trial $n-1$ prospective target and trial n target, the simulation model produced a switch-related reduction of the relationship (i.e., from .41 to .14). For the paths between $n-1$ distractor to n distractor for no-switch and switch trials, we find exactly the opposite trend (from .14 to .42). Generally, coefficients are more extreme with targets as dependent variable for no-switch trials and with distractors as dependent variable for switch trials. Further probing of the task unit activation dynamics in the model simulation explains why this is the case. The task demand unit activation values are bound between -1 and $+1$ and on no-switch trials for targets, and on switch trials for distractors, they were closer to the zero value, which is in the middle of the allowed range. Given the sigmoid activation function used to convert unit input to activation values, units with activity in this region are more sensitive to input variations than units with lower or higher activity levels. Depending on path, the effects of these activation dynamics can produce patterns that are either consistent with the carryover model, the configuration model, or neither. In the aggregate, however, the simulation results are consistent with the predictions of the carryover model. This becomes apparent when we look at RTs as dependent variable. Here, these "lower-level" influences average out, and the pattern of switch effects on coefficients is clearly consistent with the configuration model.

on the currently relevant task, these bars either served as the target or the distractor, whereas the third object served as neutral object. The three bars were presented equally spaced in three of 12 possible positions on a virtual circle (circle radius = 100 pixels or 3.4°; see Figure 1), thus constituting four possible rotations of a equilateral, virtual triangle. On each trial, stimulus positions were constrained to fall on the vertices of one of four possible equilateral triangles necessary to capture all 12 possible positions on the circle. Triangle constellations were never repeated from one trial to the next, and assignment of target, distractor, or neutral object to positions within the triangle was random. Thus, there was neither trial-to-trial predictability nor repetitions of stimulus locations.

Task cues instructed subjects to perform one of two tasks. For the "gap" task, they made a "left" response when the gap was on the bottom and a "right" response when the gap was on the top of the target object. For the "color" task, light blue meant "left" and dark blue meant "right." In Experiment 1, subjects responded by moving their eyes to left or right response regions, which began 412 pixels (13.9°) to the left or the right of the midline of the screen. Once the eyes moved into a response area, the RT was recorded, the stimuli disappeared, and the next trial began. Eye movements were used to record responses, hoping that this would encourage subjects to fixate potentially relevant objects for solving the tasks. This is a somewhat unusual aspect of our experimental setup that may have affected the generalizability of our results. However, the basic pattern of results was very similar to the one obtained in Experiment 2, in which we used regular key responses (i.e., left vs. right arrow keys).

For each trial, the relevant task was signaled by verbal task labels presented visually at the center of the screen. As explained above, the cue was selected from pairs ("gap" vs. "color", "space" vs. "hue") to avoid immediate cue repetitions. Tasks were selected randomly with the constraint that a task switch occurred with a probability of $p = .25$ in Experiment 1. For Experiment 2, task switch probability was either $p = .25$, $p = .50$, or $p = .75$ between three randomly assigned groups of subjects ($n = 24$ per group).

The interval between the start of each trial and the stimulus was 1,100 ms. Within that interval, the cue was presented for 200 ms either after 100 ms, leaving a cue–stimulus interval (CSI) of 1,000 ms, or after 800 ms, leaving a CSI of 300 ms.

Subjects were seated with their eyes approximately 50 cm from the screen and with their head supported by an adjustable chin rest. The session began with three 40-trial practice blocks during which subjects received error feedback. In Experiment 1, these were followed by 17 test blocks of 40 trials without trialwise error feedback. In Experiment 2, incentives were used to motivate subjects to use the preparatory interval. Subjects earned ¢1 for each trial with an RT that was faster than the average across all previous blocks, as long as accuracy remained at or above 90%. Subjects were not informed of this incentive scheme until after Block 4 of the experiment. During Block 4, the overall average RT was computed separately for no-switch and switch trials, and this average was updated for each new block. Subjects were instructed that no-switch and switch transitions were averaged separately in order to avoid any selective bias for the one or the other. At the end of each block, subjects were shown how much additional money they had earned (if any) along with average accuracy and RT (across no-switch and switch trials), and at the end of the experi-

ment they were paid the amount earned (usually between $3 and $4). Blocks 5–20 were used for data analysis.

Eye position registration was calibrated at the beginning of the experiment and recalibrated every four blocks. Each block's average eye positions during the presentation of the task cue were used to correct for possible drifts across the experimental session.

Analysis of Eye Movements

Eye positions were measured at a rate of 1000 Hz. Fixations were defined as the absence of either a blink or a saccade, where a saccade was identified for each pair of successive samples for which the eye's velocity surpassed 30°/s or the acceleration surpassed 8,000°/s.[2] Fixations before or after blinks were not excluded from the data—doing so would not affect the results. Around each object center, 1.7° (i.e., half the distance between cue and object) nonoverlapping circles was defined. Fixations within these circles were categorized as being directed toward that object. In the present analyses, we focused on the initial 700 ms within a trial, which are most likely to reflect attentional selection processes, rather than later rechecking. In fact, inspection of individual eye movement probability curves revealed that within these initial 700 ms, the pattern was relatively consistent across individuals, whereas later probability curves were much more idiosyncratic and varied.[3] In part, this simply reflects the fact that given average RTs of around 800 ms, data density strongly decreased beyond 700 ms.

Across all conditions, the probability of trials with target, distractor, and neutral fixations within the first 700 ms of a trial was .47 ($SD = .20$), .17 ($SD = .10$), and .06 ($SD = .06$), respectively, for Experiment 1 and .54 ($SD = .22$), .18 ($SD = .11$), and .07 ($SD = .05$) for Experiment 2. These numbers changed only slightly when looking at the entire trial duration (Experiment 1: .52, .19, .08; Experiment 2: .61, .22, .10). The fact that target fixations hovered around 50% might appear surprising at first. However, this reflects the two opposing criteria we needed to optimize for in the stimulus display. On the one hand, we needed to achieve sufficient stimulus-elicited task competition, which decreases with between-object distance. On the other hand, we needed to ensure that eye movements to the target versus the distractor are in fact indicative of the efficiency of attentional control. We could have easily enforced close to 100% eye movements by moving the objects even further apart. However, this would have turned the paradigm into a visual search task and eliminated much of the between-task competition. Obviously, our hope was that the fixations we did observe are meaningful indicators of attentional selection—overall, our results indicate that this was in fact the case.

---

[3] As we report in the context of analysis of between-trial dynamics, for the 0–700 ms phase, the correlation between target-directed fixations and RTs were negative (see Figure 7), which is consistent with the idea that these fixations indicate efficient orienting toward task-relevant information. However, when analyzing the full trial duration, this relationship reverses. Likely, this reflects both the effects of rechecking and the fact that the longer the RT, the greater the opportunity for target-related fixations, which in turn drives up the correlation between target fixations and RT. It is because of such complexities that we decided that fixations from the early trial period provide a more unambiguous reflection of attentional selection.

To reduce overall complexity of the present article, the considerable differences in overall fixations found across the two task dimensions are ignored. The color deviant generally attracted more eye movements than the gap deviant, both as a target (Experiment 1: .66 vs. .29, Experiment 2: .72 vs. .35) and distractor (Experiment 1: .21 vs. 10, Experiment 2: .24 vs. .11). However, we made sure that the general pattern of results to be reported below replicated across both task dimensions. In addition, for the analysis of across-trial dynamics, the effect of task dimension was statistically controlled for. Finally, the effect of specific value repetitions (e.g., when the color deviant carried the exact same color across trials) was also ignored, but only after ensuring that including this factor would not have affected the present conclusions.

Finally, the analyses reported above indicate that there were more than twice as many distractor fixations than fixations of the neutral stimulus, suggesting that competition during attentional selection was mainly about *potentially* task-relevant objects, and little additional information was contained in fixations to the neutral object. Therefore, only data on target and distractor fixations are reported.

Results: Within-Trial Dynamics

RTs and accuracy. We eliminated the first trial of each block, error trials, trials following an error, and all trials with RTs longer than 3,000 ms from data analysis. Figure 4 shows RTs and errors as a function of the switch and the CSI contrast for both Experiment 1 and Experiment 2. As apparent, there were substantial switch costs, as well as a reduction of costs as a function of CSI for the 25% conditions; for Experiment 2, switch costs and the preparation effect on switch costs were reduced as a function of switch rate. We used a CSI × Switch × Response Congruency analysis of variance (ANOVA) for Experiment 1—with added linear and quadratic contrasts for the three-level switch frequency factor in Experiment 2. *Response congruency* refers to whether relevant and irrelevant stimuli afford the same or different responses. To keep overall complexity at bay, we do not dwell on

this factor too much in the present analyses (however, the Qualifications section below).

Experiment 1. Both the CSI and the switch main effects were highly reliable: CSI, $F(1, 12) = 28.04$, $p < .001$, $\eta^2 = .70$; switch, $F(1, 12) = 42.79$, $p < .001$, $\eta^2 = .78$, whereas the congruency main effect was not significant, $F(1, 12) < 1.0$. The interaction between CSI and the switch factor approached reliability, $F(1, 12) = 4.64$, $p = .052$, $\eta^2 = .28$. No other effect was reliable. For errors, there was a significant switch main effect, $F(1, 12) = 19.84$, $p < .01$, $\eta^2 = .62$, as well as a main effect for response congruency, $F(1, 12) = 29.86$, $p < .001$, $\eta^2 = .71$, which was modulated by a significant Switch × Congruency interaction, $F(1, 12) = 9.18$, $p = .01$, $\eta^2 = .43$. Congruency effects (incongruent − congruent) were 1.5% on no-switch and 4.7% on switch trails.

Experiment 2. We found highly reliable CSI × Switch, $F(1, 69) = 16.19$, $p < .001$, $\eta^2 = .19$, and Linear Frequency × CSI × Switch interactions, $F(1, 69) = 25.67$, $p < .001$, $\eta^2 = .26$. The size of the preparation effect on switch costs was 76 ms, $F(1, 23) = 31.00$, $p < .001$, $\eta^2 = .57$; 35 ms, $F(1, 23) = 7.26$, $p < .05$, $\eta^2 = .21$; and −19 ms, $F(1, 23) = 2.07$, $p < .17$, $\eta^2 = .06$, for the 25%, the 50%, and the 75% frequency groups. There was no main effect for the response congruency factor, $F(1, 69) = 0.57$; however, congruency interacted with CSI, $F(1, 69) = 15.77$, $p < .01$, $\eta^2 = .19$, and there was a reliable three-way interaction between CSI, switch, and congruency, $F(1, 69) = 4.72$, $p < .05$, $\eta^2 = .6$. Congruency effects for short-CSI, no-switch, and switch trials were 24 ms, $t(71) = 2.56$, $p < .05$, and 14 ms, $t(71) = 1.75$, $p < .09$. The corresponding scores for long-CSI trials were −23 ms, $t(71) = 3.54$, $p < .01$, and 1 ms, $t(71) = 0.17$, $p > .8$. There were no other higher order interactions involving the congruency factor. The negative congruency effect for long CSIs suggests that preparation counteracted response conflict (see also Monsell & Mizon, 2006). However, relative to the rest of the literature (including Experiment 1) where typically incongruent trials are slower than congruent trials, this is a somewhat unusual result.
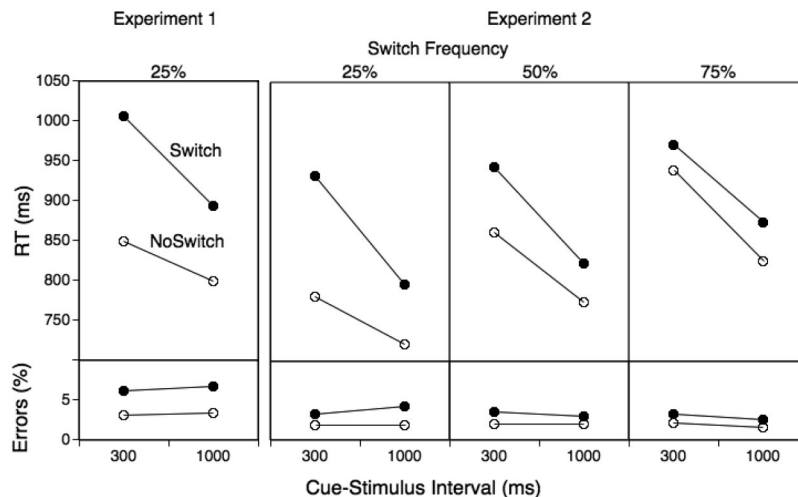

Figure 4. Average response times (RTs) and errors as a function of cue–stimulus interval (CSI) and task switch in Experiment 1 and as a function switch frequency, CSI, and task switch in Experiment 2.

For errors, there was a small, but statistically highly reliable switch main effect, $F(1, 69) = 44.66$, $p < .01$, $\eta^2 = .39$; a congruency main effect, $F(1, 69) = 40.27$, $p < .01$, $\eta^2 = .37$; and an interaction between these two factors, $F(1, 69) = 10.67$, $p < .01$, which was further modulated through the linear frequency contrast, $F(1, 69) = 15.67$, $MSE = 5.57$, $\eta^2 = .13$, $p < .01$. The switch effect on errors was larger on incongruent trials and when switch frequency was low. In Figure 4, the error switch cost appears to be a little larger with a low-switch frequency, which is in the opposite direction to the effect for RTs. However, neither the corresponding three-way interaction, $F(1, 69) < 1.3$, nor the four-way interaction including the congruency factor was even close to reliable, $F(1, 69) < .15$.

Summary.   Our RT results generally replicated the pattern reported by Monsell and Minzon (2006). The fact that the CSI × Switch interaction was not quite reliable in Experiment 1 probably reflects relatively low statistical power; the same effect in Experiment 2 was highly reliable. In Experiment 2, we also found that switch and preparation effects decreased with switch frequency, suggesting that subjects were less likely to maintain task sets across no-switch trials when frequent switches are required.

Eye Movements: Onset of Task-Congruent Attention

Figure 5 shows the proportion of all trials on which a fixation fell within the target or distractor regions (i.e., fixation probabilities) as a function of time in Experiments 1 and 2. As apparent, in all conditions, the eyes did not move until about 250–300 ms after stimulus onset. Although eventually the majority of fixations outside the cue region was directed toward the target, there was an initial phase with no differentiation between target or distractor fixations (i.e., circles vs. squares with the same color and filling). Moreover, there were large condition differences in the point in time at which probabilities of target-directed and distractor-directed fixations separate. As elaborated in the introduction, we suggest that this is the point at which top-down control over attentional orienting becomes apparent. Inspection of Figure 5 suggests that on short-CSI switch trials there were about twice as many "incorrect" eye movements to the distractor than on corresponding no-switch trials, at least when switch frequency was 50% or higher. As a result, the critical point of separation between target and distractor eye movements was particularly late in these conditions.

For a better appreciation of these effects, we computed the difference scores between the probability of target versus distractor fixations (see Figure 6). For the 25% switch frequency conditions in both experiments, these difference curves confirm that the onset of task-congruent attention occurs about 100 ms later on short-CSI switch trials than on short-CSI no-switch trials. This is consistent with the configuration model's prediction of a switch-related delay in the onset of task-congruent attention. Furthermore, this effect was eliminated for long-CSI trials, a result that is consistent with the prediction that configuration can occur in a proactive manner. In contrast, in the 75% condition of Experiment 2, a qualitatively different pattern emerged. Here, the onset of control was delayed not only for switch but also for no-switch trials in the short-, but not in the long-CSI condition. This result confirms that configuration may be mandatory on switch trials, but can occur on no-switch trials when there is little incentive for
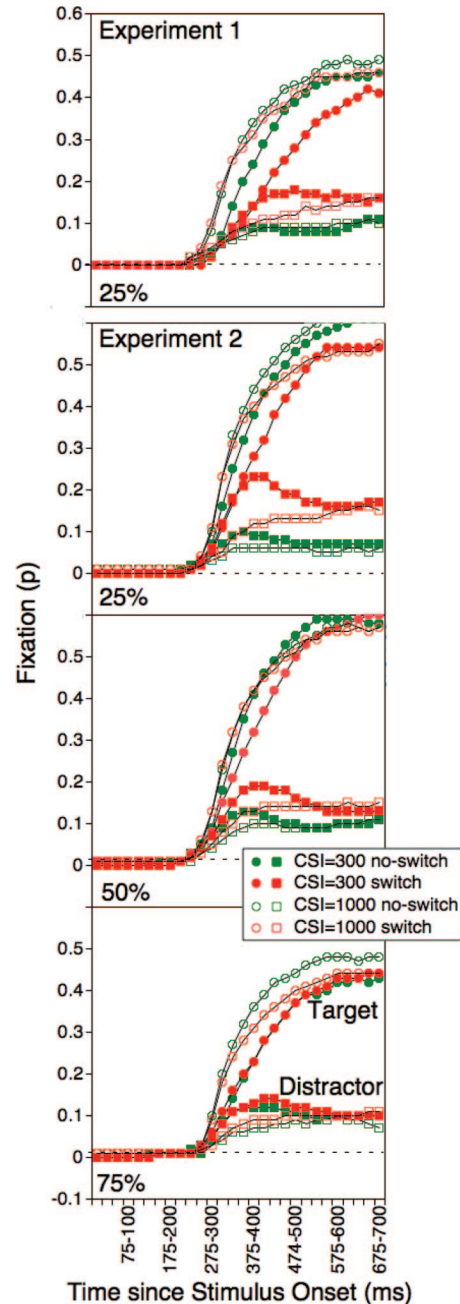

Figure 5. Probability of target and distractor fixations as a function of time, cue–stimulus interval (CSI), and task switch in Experiment 1, and as a function of time, switch frequency, switch, and CSI in Experiment 2.

maintaining the current task set (i.e., because of high-switch frequency). Finally, the 50% condition is similar to the 25% condition. The delay in control on short-CSI switch trials may not be as crisp; however, the overall trajectory is clearly shifted as a function of the task switch.
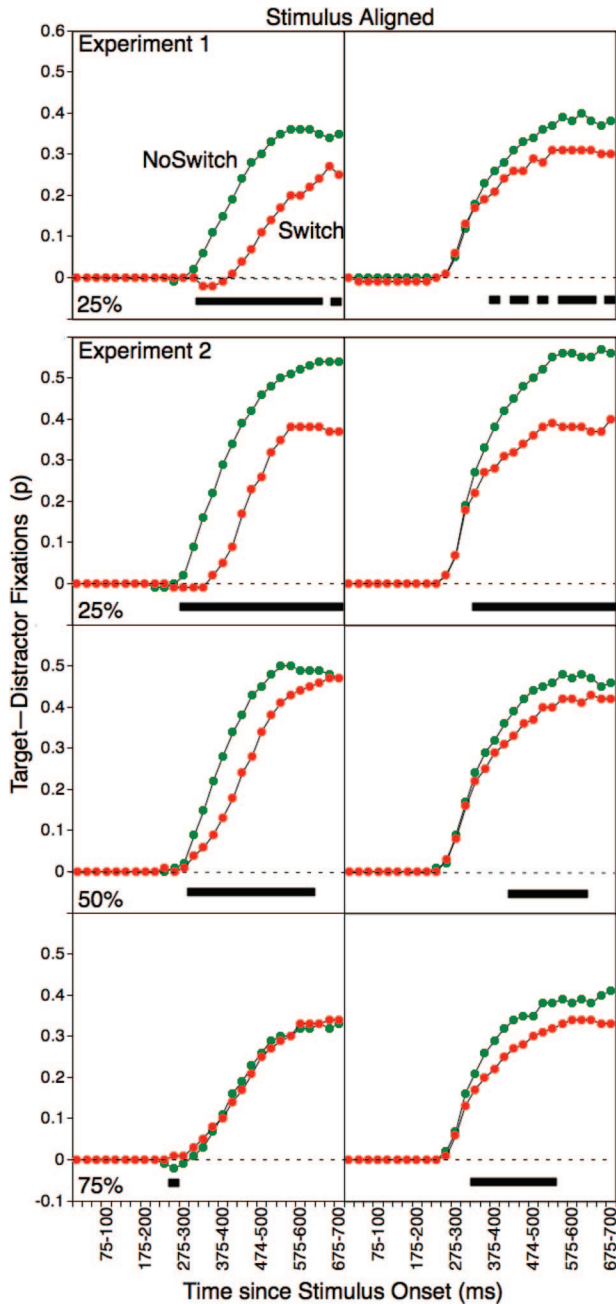
To confirm these observations statistically, we determined for each subject and condition the first 25-ms bin during which the target-minus-distractor difference curves (see Figure 6) reached 4% or more for at least two successive bins. The onset latencies for

both experiments are presented in Table 1. For Experiment 1, The CSI factor, $F(1, 12) = 37.44$, $p < .001$, $\eta^2 = .76$; the switch factor, $F(1, 12) = 11.68$, $p < .01$, $\eta^2 = .49$; and the CSI × Switch interaction, $F(1, 12) = 17.43$, $MSE = 3094.95$, $p < .03$, $\eta^2 = .59$, were all reliable. For Experiment 2, one subject in the 25% frequency group and three subjects in the 75% condition never reached the criterion for determining the top-down control onset in

Table 1

Average Time Points (and Standard Deviations) at Which Differences Between Individual Eye Movement Probability Functions Toward the Target and the Distractor Reached at Least 4% for Two Consecutive 25-Millisecond Time Windows as a Function of CSI, and the Switch Contrast for Experiment 1 and as a Function of Switch Frequency, CSI, and the Switch Contrast for Experiment 2

| | CSI | | | |
| | Short | | Long | |
Switch frequency	No switch	Switch	No switch	Switch
Experiment 1				
25%	329 (37)	456 (132)	290 (30)	288 (44)
Experiment 2				
25%	302 (19)	396 (60)	274 (18)	286 (27)
50%	306 (48)	362 (73)	278 (36)	278 (32)
75%	361 (56)	390 (151)	285 (41)	305 (85)

Note. CSI = cue–stimulus interval.

at least one of the conditions. These subjects, who generally showed few eye movements, were eliminated from the analysis. The ANOVA of onset latencies revealed a linear Frequency × CSI × Switch interaction, $F(1, 65) = 14.79$, $p < .01$, $\eta^2 = .19$. Highly reliable Switch × CSI interactions were obtained in the 25% and the 50% frequency conditions, $F(1, 22) = 53.57$; and, $F(1, 23) = 17.86$. However, for the 75% condition, neither the switch effect, $F(1, 20) = 1.02$, nor the CSI × Switch interaction, $F(1, 20) = 1.02$, was reliable. The only reliable effect was a highly significant delay of the onset latencies for short versus long CSIs, $F(1, 20) = 42.84$. Finally, across all frequency conditions, the switch effect was eliminated for long CSIs (all $Fs < 1.4$, $p > .2$).[4]

Eye Movements: Postonset Selectivity

So far, our results indicate that the onset of control is delayed for short-CSI switch trials when switch frequency is low to moderate. This result is consistent with the idea of a discrete processing step that needs to be completed prior to the onset of task-congruent attention. Now we turn to potential switch-related effects on the eventually adopted attentional settings. In order not to confound contributions of delayed onsets of top-down control and eventual later effects, it is useful to look at the long-CSI condition, for which there was no substantial switch cost in the onset of task-

Figure 6. The probability of target minus the probability of distractor fixations as a function of time, cue–stimulus interval (CSI), and task switch in Experiment 1 and as a function of time, switch frequency, CSI, and switch in Experiment 2. The black bars at the bottom indicate each 25-ms bin with significant differences between no-switch and switch trials ($p < .05$).

---

[4] One possibility is that the frequency-specific modulation of switch effects on eye movement probability curves is not so much a result of the global selection context, but rather of local sequential effects: In case of low switch frequency, switch trials rarely follow other switch trials, whereas the reverse is the case when switch frequency is high. However, we also looked at eye movement probability curves across all frequency conditions conditional on whether the previous trial was a switch or a no-switch trial (i.e., keeping local context constant). In terms of the most critical qualitative characteristics, probability curves were very similar across all frequency conditions, in particular for the onset-specific effects, suggesting that it is mainly the global context that drives the frequency-specific modulation seen in Figures 5 and 6.

congruent attention. As Figures 5 and 6 show, task-congruent selectivity was greater on no-switch than on switch trials, even on long-CSI trials. Figure 6 also shows the results of simple *t* tests comparing switch with no-switch scores for each 25-ms bin. For both experiments, there was a substantial portion of bins with significantly fewer task-congruent eye movements in switch than in no-switch trials. Thus, even though there was no difference in the onset of top-down control between switch and no-switch trials on long-CSI trials, the attentional settings were not sufficient to ward off the persistent interference from the competing task. As mentioned in the introduction, such a persistent switch-related effect on task-congruent attention is difficult to reconcile with a strict interpretation of the compound-cue model (Schneider & Logan, 2005).

An interesting exception to this pattern occured in the 75% short-CSI condition (Experiment 2). There was not only no switch effect on the onset of attentional selectivity, but even beyond the onset no switch effect emerged (except for an early reverse effect in the direction of greater selectivity for switch trials). It is interesting that this absence of task-congruent selectivity for short-CSI trials gives way to strong selectivity for long-CSI trials. Given that short- and long-CSI trials were randomly intermixed suggests that attentional selectivity can be flexibly changed on a trial-by-trial manner.

Discussion: Within-Trial Dynamics

The discrete shift in onset of task-congruent attention is consistent with the configuration model, and it is difficult to explain from the perspective of the compound-cue model (Schneider & Logan, 2005). However, the situation is a bit more complicated for the Gilbert and Shallice (2002) model, which can account for a discrete shift in onset of task-congruent processing. Carry-over of activation from the preceding trial on switch trials leads to initial task demand unit activation that favors the currently irrelevant task (e.g., see Figure 5 in Gilbert & Shallice, 2002). Only after this reversal is overcome can task-congruent selectivity begin. Even though an onset delay is not inconsistent with the task-level selection model, the fact that this effect is contingent on the overall switch frequency is more problematic for the Gilbert and Shallice model. After all, in this model, switch costs are the result of passive priming of task demand node activation, with no mechanism in place that allows for context-specific adaptations of such priming effects.

Even if we grant that the carryover model could be amended to allow context-specific adaptations, there is one aspect in our data that such an "upgraded" carryover model cannot account for: Although we found substantial switch costs in eye movement parameters in case of long preparatory intervals for the 75% condition, such effects were absent for the short-CSI condition. From the perspective of the configuration model, it is plausible that on short-CSI trials, the absence of the cue within the fixed response-to-stimulus interval is used by subjects as a trigger to abandon the last relevant task and turn toward the alternate task, resulting in a lack of switch effects on attentional selectivity. In contrast, on long-CSI trials, the cue can be used to flexibly engage either in task maintenance (thereby benefitting from the attentional setting carried over from the preceding trial) on no-switch trials or

in abandoning the previous and retrieving the next task on switch trials.

It is difficult to see how the carryover model (Gilbert & Shallice, 2002) would explain such a pattern. After all, the main point of this model is that one and the same parameter constellation is used across trials, so that no "special" executive operation needs to come in and tweak processing on particular trials. Obviously, any parameter constellation that produces no switch costs on short-CSI trials would do the same on long-CSI trials—which is clearly inconsistent with the results we obtained (see Figure 6).

To conclude, the switch-specific delay in the onset of control is consistent with the configuration model and can rule out some (e.g., Schneider & Logan, 2005), but not all alternative accounts. Evidence from the switch frequency manipulations adds empirical constraints that are easily handled by the configuration model, but difficult to accommodate by the task-level selection account (Gilbert & Shallice, 2002). In the next section, we turn to the analysis of trial-to-trial dynamics to provide an additional source of information for distinguishing between competing models of control.

Results: Trial-to-Trial Coupling Dynamics

We now turn to eye movement analyses of the coupling between control settings across consecutive trials. In the introduction, we had derived competing predictions about across-trial coupling dynamics from the carryover model and the configuration model (see Figures 2 and 3), which we now put to an empirical test using mixed linear model analyses. Recall that this empirical test requires two complementary specifications. The first specifies eye movements on trial $n-1$ in terms of what is relevant on that trial and is ideal for testing the task-level selection model (see Figure 2a and top-left panel of Figure 3). It is presented here in terms of Models 1a, 2a, and 3a. The second specifies eye movements on trial $n-1$ with relation to what is target and distractor on trial n (i.e., as prospective targets/distractors) and allows a direct test of the configuration model (see Figure 2b and bottom-right panel of Figure 3). These specifications are presented as Models 1b, 2b, and 3b. In Models 1a and 1b, the dependent variable was whether or not the eyes moved toward the target (coded as off-target = 0 vs. on-target = 1) during the first 700 ms of a trial; in Models 2a and 2b, the corresponding analyses for eye movements to the distractor were used. In each of these models, the pre-700 ms, trial $n-1$ eye movements directed to the target or the distractor (again coded as 0 vs. 1), switch (coded as no-switch = 0 and switch = 1), and Target × Switch and the Distractor × Switch interactions served as fixed effect predictors. Note that given that we model here the probability of eye movements to either the target or the distractor, unstandardized coefficients can be interpreted in terms of probabilities. For example, a coefficient of .09 between trial $n-1$ target and trial n target on switch trials (see Figure 7a) indicates that the probability of the eyes going toward the target increases by $p = .09$ if the eyes were on the target on the previous trial.

Trial n RTs were used in Models 3a and 3b as the dependent variable and both trial n and trial $n-1$ eye movements to targets and distractors as predictors along with the experimental variables. Given that here the dependent variable is RTs, the coefficients can be interpreted in terms of milliseconds. For example, the coefficient for the trial n target predictor is −86 on no-switch trials (see
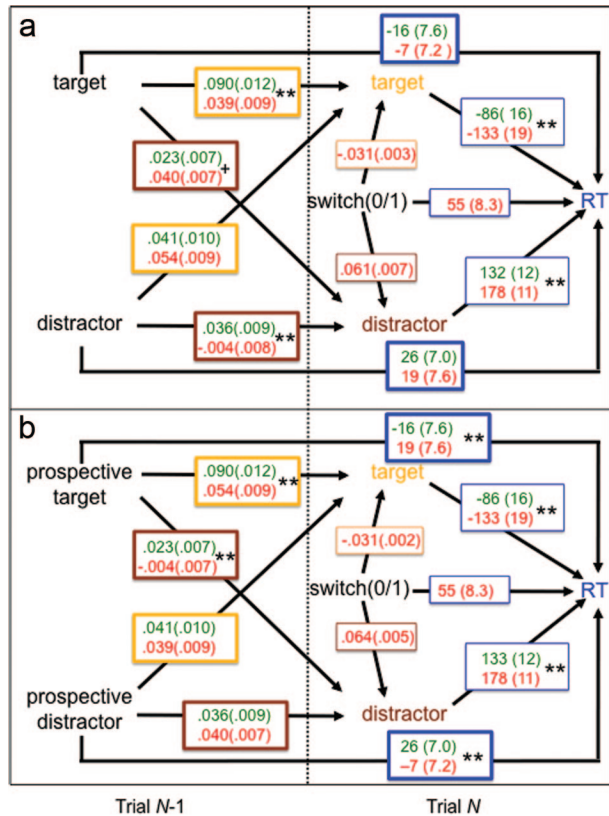
Figure 7. Results of the two complementary sets of analyses to test the task-level selection model (corresponding to regression Models 1a, 2a, 3a; top panels of Figure 3) and the configuration model (corresponding to regression Models 1b, 2b, 3b; bottom panels of Figure 3). Coefficients and dependent variables associated with the same model are contained in boxes with the same color (Model 1: orange; Model 2: brown; Model 3: blue). For each of the models, unstandardized coefficients and the corresponding standard errors were estimated through separate analyses for no-switch and switch trials, whereas the difference of the coefficients between no-switch and switch trials was tested through the relevant interaction terms in the combined analyses. The coefficients for the switch effects were also taken from the combined analyses. Coefficients linked to eye movements as dependent variables represent fixation probabilities, whereas coefficients linked to response times (RTs) as dependent variable are in millisecond units. Coefficients that are about twice the size of the standard error are significantly different from zero. Symbols indicate approximate significant differences between no-switch and switch coefficients ($^+$ $p < .1.$ ** $p < .01.$). Boxes with bold outlines indicate coefficients that are critical for testing the two alternative models of control dynamics.

Figure 7a), suggesting that an eye movement toward the target reduces same-trial RTs, on average, by that amount.

For all models, we removed linear and quadratic effects of blocks and trials, as well as task effects in an initial preprocessing step. Intercepts and all main effects were specified as subject-level random effects and thus could vary across subjects. This ensures that the fixed effects we report represent the averages of relationships that were estimated within subjects. Models were analyzed in R (R Development Core Team, 2008), using the lmer package (Bates et al., 2008).

The fixed-effect results are presented in Figure 7 and in Table 2. To foreshadow our findings, the coupling between trial $n-1$ and trial n eye movements was somewhat complex and favored the task-level selection account, but it was also relatively inconsequential for the prediction of RTs through previous-trial eye movements. In contrast, the coupling pattern between trial $n-1$ eye movements and trial n RTs strongly confirmed the predictions of the configuration model. We begin by presenting the coupling between trial $n-1$ and trial n eye movements (i.e., Models 1a, 2a, 1b, and 2b).

"Eye-to-Eye" Coupling Dynamics

In order to test the task-level selection model, we begin with the specification in terms of task relevance (i.e., see top-left panel of Figure 3). If the model is correct, coefficients for "horizontal" ($n-1$ target to n target or $n-1$ distractor to n distractor) should be positive, and "diagonal" coefficients ($n-1$ target to n distractor or $n-1$ distractor to n target; see top-left panel of Figure 3) should be negative. Most importantly, on switch trials, positive coefficients should move into the negative direction and negative coefficients in the positive direction. Consistent with expectations, we actually did find highly reliable positive, horizontal relationships (see, e.g., Figure 7a) for no-switch trials. For example, the probability of moving the eyes to the trial n target increased by $p = .09$ if the eyes were on the trial $n-1$ target. However, the diagonal relationships were positive instead of the predicted negative relationships. Turning to switch trials, coefficients changed significantly in the direction predicted by the task-level control model (e.g., trial $n-1$ target to trial n target from .09 for no-switch to .039 for switch trials), except for the trial $n-1$ distractor to trial n target relationship, which was significantly positive and did not differ between no-switch and switch trials (.041 vs. .054). Thus, for three out of four coefficients, the switch factor interacted with the trial $n-1$ predictors in the expected manner. At the same time, when the model predicted a flip of coefficients into negative territory, coefficients remained positive, or at best near zero. One possibility is that switch-specific effects sit on top of an unspecific positive, trial-to-trial correlation of a general tendency to move the eyes to any object on the screen.

It is noteworthy that predictions of the task-level selection model were violated most blatantly for the relationships between trial $n-1$ distractor and trial n target fixations. Not only were they positive rather than negative on no-switch trials (no-switch = .041, switch = .054), they were also not reliably affected by the switch contrast. Interestingly, such a pattern is consistent with a conflict monitoring mechanism (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001) where experience of conflict on trial $n-1$ (i.e., indicated through eye movements to the distractor) tightens control and therefore increases target-directed processing on trial n. Interestingly, the fact that this relationship was not modulated by the switch factor suggests that conflict-triggered control might occur in a way that is more consistent with general than task-specific control (see, e.g., Freitas, Bahar, Yang, & Banair, 2007; but see Goschke, 2000).

Model Variants 1b and 2b allow direct tests of the configuration model by incorporating trial $n-1$ prospective targets and distractors (i.e., lower right panel in Figure 3). For no-switch trials, results are identical to Models 1a and 2a. However, for switch trials, the configuration model, but not the carryover model, predicts that "horizontal" relations (e.g., prospective target $->$ trial n

Table 2
Fixed-Effect Results for Models Presented in Figure 7

Effect	Unstandard coeff.	SE	t
Model 1a: Trial n target as dependent variable			
intercept	**0.011**	**0.003**	**4.38**
$n-1$ target	**0.089**	**0.010**	**8.83**
$n-1$ distractor	**0.040**	**0.009**	**4.74**
switch	**−0.018**	**0.004**	**−4.86**
$n-1$ Target × Switch	**−0.050**	**0.010**	**−4.80**
$n-1$ Distractor × Switch	0.014	0.011	1.29
Model 1b: Trial n target as dependent variable			
intercept	**−0.011**	**0.003**	**4.41**
$n-1$ prosp target	**0.090**	**0.010**	**8.97**
$n-1$ prosp distractor	**0.040**	**0.009**	**4.60**
switch	**−0.018**	**0.004**	**−4.89**
$n-1$ Prosp Target × Switch	**−0.034**	**0.011**	**−3.07**
$n-1$ Prosp Dist. × Switch	−0.002	0.010	−0.02
Model 2a: Trial n distractor as dependent variable			
intercept	**−0.032**	**0.002**	**−13.36**
$n-1$ target	**0.023**	**0.007**	**3.51**
$n-1$ distractor	**0.033**	**0.008**	**4.22**
switch	**0.056**	**0.003**	**15.99**
$n-1$ Target × Switch	0.018	0.009	1.96
$n-1$ Distractor × Switch	**−0.032**	**0.010**	**−3.25**
Model 2b: Trial n distractor as dependent variable			
intercept	**−0.032**	**0.002**	**−13.37**
$n-1$ prosp target	**0.023**	**0.007**	**3.17**
$n-1$ prosp distractor	**0.032**	**0.008**	**4.18**
switch	**0.056**	**0.004**	**15.98**
$n-1$ Prosp Target × Switch	**−0.023**	**0.010**	**−2.38**
$n-1$ Prosp Dist. × Switch	0.009	0.010	−0.98
Model 3a: Trial n RT as dependent variable			
intercept	**−64.810**	**9.601**	**−6.750**
n target	**−77.819**	**16.935**	**−4.595**
n distractor	**142.004**	**10.918**	**13.007**
$n-1$ target	**−16.994**	**6.990**	**−2.431**
$n-1$ distractor	**29.519**	**7.115**	**4.008**
switch	**54.726**	**8.223**	**6.655**
n Target × Switch	**−50.105**	**9.602**	**−5.218**
n Distractor × Switch	**38.329**	**10.0931**	**3.798**
$n-1$ Target × Switch	7.468	8.938	0.836
$n-1$ Distractor × Switch	−10.189	9.652	−1.056
Model 3b: Trial n RT as dependent variable			
intercept	**−64.977**	**9.673**	**−6.717**
n target	**−77.866**	**16.935**	**−4.598**
n distractor	**142.213**	**10.939**	**13.001**
$n-1$ target	**−16.685**	**6.737**	**−2.477**
$n-1$ distractor	**28.826**	**7.224**	**3.990**
switch	**54.875**	**8.251**	**6.649**
n Target × Switch	**−50.681**	**9.605**	**−5.213**
n Distractor × Switch	**38.336**	**10.094**	**3.789**
$n-1$ Target × Switch	**36.002**	**9.355**	**3.848**
$n-1$ Distractor × Switch	**−37.765**	**9.218**	**−4.097**

Note. Model 1a: Trial n target regressed on trial $n-1$ distractor and target, switch contrast, as well as all relevant interactions. Model 2a: Trial n distractor regressed on trial $n-1$ distractor and target, switch contrast, as well as all relevant interactions. Model 3a: RTs regressed on trial n distractor and target, trial $n-1$ target and distactor, switch contrast, and all relevant interactions. Models 1b, 2b, 3b used trial $n-1$ prospective distractor and targets instead of trial $n-1$ targets and distractors. All "target" and "distractor" variables are binary and represent whether or not the eyes moved to either the target or the distractor on that trial. For Models 1a, 1b, 2a, and 2b, coefficients represent probabilities; for Models 3a and 3b, coefficients represent response times in milliseconds. Variables were residualized with regard to linear and quadratic effects of block and trial as well as task differences before being submitted to these analyses. Each model also included as random effects the intercept as well as all main effects. Coefficients printed in bold were significant. The criterion for a $p = .05$ significance level is a coefficient close to two standard errors (i.e., $ts > 1.9$). The degrees of freedoms for t values are not known exactly for linear mixed models. However, given the large number of observations in our analyses, the t distribution has converged to the standard normal distribution. In this case the 2-*SE* criterion is close to the conventional two-tailed 5% level of significance (e.g., Baayen et al., 2008, Note 1). Note that small numerical differences in coefficients between this table and Figure 7 are due to the fact that for the figure, coefficients and standard errors for no-switch and switch conditions were estimated through separate analyses. coeff. = coefficient; prosp = prosepective; dist. = distractor.

target) should turn negative on switch trials, whereas if anything, diagonal relations should become positive. For example, if the eyes were on the same trial $n-1$ object type that will become distractor on trial n (i.e., the prospective distractor), then a switch-related reconfiguration should eliminate or even reverse any cross-trial coupling that may exist on no-switch trials. However, as we see in Figure 7b for this specific example, the probability of an eye movement toward the distractor increases by $p = .036$ if the eyes were on the trial $n-1$ prospective distractor, and if anything, this probability increases to $p = .04$ on switch trials. Only the coefficients between the prospective target and the trial n target behave in a manner that is consistent with the configuration model. The three other coefficients either move in the reverse direction or are unaffected by the switch factor. Thus, for eye movements as dependent variable, trial-to-trial dynamics are more consistent with the task-level selection than the configuration model. However, overall, the positive evidence in favor of the carryover model is weak and potentially clouded by additional factors, such as conflict-triggered recruitment of control.

"Eye-to-RT" Coupling Dynamics

With Model 3a, we test whether the influence of trial $n-1$ eye movements on trial n RTs is consistent with the task-congruent carryover model (i.e., top-left panel of Figure 3). We also enter here trial n eye movements as predictors, thus controlling for the information represented on the level of trial n eye movements (i.e., the information captured with Models 1a and 2a). Thus, now coefficients represent the average RT benefit or cost as a function of trial $n-1$ eye movements. Given the inclusion of trial n eye movements, we can also estimate the within-trial n predictive relations between eye movements and RT. As shown in Figure 7a, for no-switch trials, target-directed fixations produced a benefit of nearly 90 ms and distractor-directed fixations a cost of about 130 ms; for switch trials, these values were even larger. This confirms that eye movements to targets and distractors actually serve as valid indicators of task-congruent attentional processing (but see Footnote 3).

Turning now to the theoretically critical trial-to-trial relations, we ask again to what extent direct paths between eye movements to trial $n-1$ targets/distractors and RTs are consistent with either the task-level selection or the configuration model. As both models predict, for no-switch trials the influence from the $n-1$ target was negative and that from the $n-1$ distractor was positive: An eye movement to the trial $n-1$ target reduced the next-trial RT by 16 ms, whereas an eye movement to the distractor increased next-trial RTs by 26 ms (see Figures 7a and 7b). For the switch trials, the carry-over model predicts that these coefficients flip signs (i.e., the trial $n-1$ target becomes the distractor, and the trial $n-1$ distractor becomes the new target; see Figure 7a). In contrast, this relationship was not reliably affected by the switch factor. Accordingly, a chi-square test comparing models with and without switch interactions for trial $n-1$ targets and distractors revealed no significant difference, $\chi^2(2) = 1.97, p > .3$.

With Model 3b, we test for an effect of switch-specific cognitive control on the relationship between prospective targets or distractors and RTs. The configuration model predicts that the cross-trial relationships flip in sign between no-switch and switch trials (see bottom-right panel of Figure 3). This is exactly what we found: If

the eyes were on the trial $n-1$ dimension that becomes target on trial n, RTs were reliably faster by 16 ms for no-switch trials, but actually reliably slower by 19 ms for switch trials. Equally the coefficients for the prospective distractors turned from reliably positive to negative (i.e., from 26 to -7, although here the negative coefficient was not reliable; see Figure 7b). The chi-square test comparing models with and without switch interactions with trial $n-1$ targets or distractors showed a highly reliable difference, $\chi^2(2) = 35.48, p < .001$. Thus, for task switches, an additional process eliminates the trial-to-trial coupling observed on no-switch trials. As elaborated in the introduction, a significant reversal for the relationship between prospective trial $n-1$ targets and RTs is consistent with the idea that active inhibition of the no longer relevant task is part of the configuration process (e.g., Mayr & Keele, 2000).

By using both eye movements and RTs as dependent variables (Models 1 and 2 vs. Model 3), we could—as a first approximation—separate between dynamics on the level of attentional selection and the level of response selection. However, it is also useful to examine the net effect of the activity on these two levels. For this purpose, we respecified Models 3a and 3b without controlling for trial n eye movement predictors. For this revised Model 3a, we found that the coefficients from trial $n-1$ target to trial n RT were reduced from -20 ($SE = 8.0$) to -5 ($SE = 7.6$), for no-switch versus switch trials, Switch × Trial $n-1$ Target: $t = 1.33$, and those for $n-1$ distractors from .29 ($SE = 7.0$) to .16 ($SE = 7.9$), Switch × Trial $n-1$ Target: $t = 1.73$.[5] The direction of these numerical, switch-related changes was consistent with the task-level control model; however, in neither case were they significant. For the revised Model 3b, we found for prospective targets a highly significant inversion of the coefficient from -20 ($SE = 8.0$) to 16 ($SE = 7.9$), and for prospective distractors, again a highly significant inversion from 29 ($SE = 7.0$) to -4 ($SE = 7.6$). Thus, even though on the level of eye movements there was some evidence for task-level selection, the effect on RTs was muted and not significant. The overall pattern was again clearly consistent with the configuration model.

Examining the coefficients in Figure 7b, it is not surprising that trial $n-1$ eye movements had only minor effects on trial n RTs. Consider for example the strongest eye-to-eye coupling: A probability of .09 by which a fixation to the prospective target increases the probability of a trial n fixation of the target on no-switch trials would translate only into a 8-ms benefit on the level of RTs (i.e., $.09 \times -87$ ms $= 7.8$ ms).

Finally, if the coupling effects reported here are indicative of a configuration model, then it would be plausible to assume that they are modulated through opportunity for proactive control or strategic settings. Therefore, we added CSI (coded as $-.5$ and $+.5$) and switch frequency (linear and quadratic contrast) as additional predictors, including all relevant interactions, which yielded a very complex model, with a total of 59 fixed-effect predictors. The pattern of theoretically critical coefficients reported above re-

---

[5] Degrees of freedom for t values are not known exactly for linear mixed models. However, with the large number of observations, the t distribution converges to the standard normal distribution and a 2-SE criterion is close to the conventional two-tailed 5% level of significance (e.g., Baayen, Davidson, & Bates, 2008, Note 1).

mained virtually unchanged, and none of the interactions including these effects were reliable. However, some interesting numerical trends emerged. Specifically, the coupling between the prospective target and RTs for both no-switch trials and for switch trials was generally reduced as switch frequency increased (i.e., the coupling for no-switch trials became more positive, that for switch trials more negative) (both effects $t > 1.1$). Thus, as switch frequency increases, the trial-to-trial coupling may be strategically reduced. We also found that both the negative coupling between the prospective target and RTs during no-switch transitions and that between the prospective distractor and RTs during switch transitions became stronger for longer CSIs ($t = 1.56$ and $t = 1.69$), indicating that trial-to-trial coupling may be at least partly dependent on proactive control processes. These are weak effects that we report here mainly for sake of completeness. However, they suggest that the existence of strategic and proactive modulations of trial-to-trial coupling effects cannot be ruled out on the basis of the present results, and tests with greater statistical power will be necessary for definitive conclusions.

General Discussion

The analysis of within-trial and across-trial eye movement dynamics provided a number of novel results about the nature of flexible task control. In the preceding sections, we have already addressed many of the specific theoretical implications. Here, we summarize and expand on the most important findings.

Within-Trial Control Dynamics

Switch-related discrete delay in the onset of task-congruent attention. When switch probability was low or moderate, the onset of task-congruent attention reflected in eye movement probability curves was postponed by about 100 ms on switch trials, but only when there was insufficient time to prepare (see Figures 5 and 6). This pattern is consistent with a discrete configuration step that is required on switch trials, that needs to precede task-congruent processing, and that can be executed proactively. In contrast, such a pattern is inconsistent with the compound-cuing model (Schneider & Logan, 2005).

Delayed onset of task-congruent attention for no-switch trials when switch frequency is high. By itself, the switch-related delay could also be consistent with the task-level selection account (see, e.g., Gilbert & Shallice, 2002). However, the switch-related delay was also strongly affected by the switch frequency manipulation. The results suggest that for high switch frequency, the delay was present even on no-switch trials (see Figure 6). This is consistent with a discrete configuration process that is under strategic control. Specifically, when switch frequency is high, subjects operate under a general change-oriented control setting in which reconfiguration and/or inhibition of the previous task tends to be initiated even on no-switch trials (see also Mayr, 2006).

CSI-dependent modulation of task-congruent attention. A rather subtle but theoretically important finding is that no-switch effects on task-congruent selectivity were apparent in eye movements when the switch frequency was high and CSI was short, whereas substantial effects emerged when the CSI was long (see Figure 6). Most likely this reflects the fact that for high switch frequency and short CSI, subjects tended to abandon previous task

settings even on no-switch trials. However, cue-driven foreknowledge on long-CSI trials replaced the a priori expectation and allowed subjects to maintain the previous trial settings on no-switch trials. Critically, given that CSI varied randomly from trial to trial, this pattern suggests a kind of online modulation of task-congruent attentional selectivity that would be difficult to account for on the basis of a passive carryover account.

Trial-to-Trial Control Dynamics

Models that explain control phenomena without invoking higher level task-unspecific control operations have been successful in predicting mean-level effects (e.g., Gilbert & Shallice, 2002; Schneider & Logan, 2005). At the same time, it has been difficult to identify unambiguous empirical markers that could be linked to task-unspecific control operations (e.g., Gilbert, 2005; Kiesel et al., 2010; Schneider & Logan, 2005). However, as we have shown here, task-level selection and configuration models make distinct predictions about how the efficiency of control settings is coupled across trials. Consistent with predictions from the task-level selection model (Gilbert & Shallice, 2002), we found that attentional settings on trial $n-1$ were predictive of attentional settings on trial n as long as tasks stayed the same across trials. In case of switch transitions, there was some evidence for a carryover pattern when looking at eye movements as dependent variable. However, for RTs as dependent variable, task-specific carryover was counteracted during switch transitions, a result that is consistent with the configuration model.

The fact that the eye-to-eye coupling pattern was different from that for eye-to-RT relationships requires some additional comments. At least at first approximation, this suggests a separation between coupling effects within the attentional/eye movement system and those that arise as at the interface between attention and response selection. A critical process in this regard may be the linking of a specific stimulus dimension to response selection (e.g., Meiran, 2000). If a previously attended object type becomes less available for being linked to response selection, this could explain why attending the trial $n-1$ prospective target leads to longer RTs and why attending the trial $n-1$ prospective distractor tends to lead to faster RTs.

Given that the task-level selection account received at least some empirical support, it would be premature to dismiss it. In fact, there is no principled reason why control could not occur both on task-specific and a more general level (e.g., Brown et al., 2007; Rougier & O'Reilley, 2002). Moreover, our results suggest as an interesting hypothesis for future research that in such a hybrid control model, task-specific carryover characterizes control dynamics within the attentional system, whereas the task-unspecific configuration is better suited to explain effects on response selection.

Another interesting aspect about the across-trial coupling we obtained (i.e., the complete flip in signs of coefficients) is that it is qualitatively consistent with the coupling dynamics produced by a model in which both activation of the relevant and inhibition of the irrelevant task is used (for complementary evidence, see also Koch, Gade, Schuch, & Philipp, 2010; Mayr & Keele, 2000). Of course, this does not rule out the possibility that alternative models could be constructed that can produce such a dynamic pattern without inhibitory processes. In fact, we conducted additional simulations, which indicated that a similar dynamic coupling pat-

tern is also produced by a hybrid model with both task-level and task-unspecific carryover of control settings. Thus, further examinations of the relative contributions of task-specific, task-unspecific, and inhibitory processes to the across-trial coupling patterns remain an important goal.

These open questions notwithstanding, our results show for the first time that information about the dynamic coupling across trials can be useful for constraining models of control. In future work, it will also be important to examine how such information can be used to characterize differences in control architecture across selection situations or individuals. In this context, it is also noteworthy that our logic of distinguishing configuration and carryover models is analogous to the successful use of dynamic covariation patterns for distinguishing between-hierarchical and nonhierarchical models of rhythm production (e.g., Krampe, Kliegl, Mayr, Engberg, & Vorberg, 2000).

What Is the Source of Interference During Task Selection?

Looking at within-trial and between-trial dynamics combined, our results seem to represent an interesting paradox. The within-trial dynamics showed that interference from the currently irrelevant task is generally increased on switch trials. However, our analyses of between-trial dynamics revealed little evidence for carryover effects (at least with regard to RTs) on switch trials. Thus, the question is: If it is not carried over from the immediately preceding trials, where exactly does the switch-related interference come from that we see in the analysis of within-trial dynamics?

It is intuitively appealing to give the immediately preceding trial and its influence on current-trial processing a special status in task-switching situations (e.g., Gilbert & Shallice, 2002; Yeung & Monsell, 2003). However, this may be misleading. There is increasing evidence that a major source of interference on switch trials arises from long-term memory (LTM) traces of selection episodes beyond the most recent trial (e.g., Mayr, 2009; Mayr & Bryck, 2005; Waszak, Hommel, & Allport, 2003). Interestingly, people also seem to inhibit the previous trial task set while switching to a new task (e.g., Mayr & Keele, 2000; Koch et al., 2010), thereby dampening the influence from the most recent past.

The question remains why processing on switch trials is particularly vulnerable to interference from past selection instances. Bryck and Mayr (2008) have argued that on switch trials, people are forced to retrieve the now relevant task rules from LTM, and it is this process of retrieval that opens up working memory—not only to relevant but also to unwanted LTM traces (see also O'Reilley, 2006). Consistent with this claim, these authors showed that an empirical phenomenon often considered a signature of task set carryover, the switch-cost asymmetry (Gilbert & Shallice, 2002), is not confined to switch trials. Rather, an asymmetry in selection costs also appears on those no-switch trials during which a retrieval attempt is likely (e.g., because of a long intertrial delay).

Qualifications

We focused here on testing two possible "baseline models": the compound-cue model (Logan & Bundesen, 2003; Schneider &

Logan, 2005) and Gilbert and Shallice's (2002) carryover model. The virtue of these two models is their simplicity and that we can derive from each clear predictions about within-trial and/or trial-to-trial dynamics. Admittedly, these tests were also slightly unfair as neither of the models was originally designed to account for eye movement dynamics. It remains to be seen how difficult it would be to amend them in order to account for the type of results presented here. Connectionist models by Brown et al. (2007) and Rougier and O'Reilley (2002) incorporate carryover of attentional settings alongside switch-specific and/or conflict-specific control processes. They may be better suited to reproduce the control dynamics reported here.[6] However, the explicit goal of both the Gilbert and Shallice model (see also Gilbert, 2005) and the compound-cue model (e.g., Schneider & Logan, 2005) has been to challenge the idea of switch-specific, high-level control processes. Therefore, we believe they are both worthy and tractable targets for this first attempt of using eye movement information to constrain models of control.

Another important limitation is that by tracking eye movements, we are assessing mainly task-switching effects on attentional input selection, but at least at first sight have little to say about the role of response selection. This is potentially problematic because traditionally, conflict during response selection has been viewed as a major source of task-switching deficits (e.g., Schuch & Koch, 2003). To determine how response selection and eye movements are related in our paradigm, we conducted additional analyses, which, for space reasons, we can present here only in summary form. First, we looked at the probability curves in a response-aligned manner. The general logic, which we borrowed from the analysis of response-aligned event-related response waves, is that if major task-switching effects arise beyond the attentional selection stage, we should see switch or preparation effects on these response-aligned functions. However, we found that effects on response-aligned curves were generally much smaller than the corresponding effects on stimulus-aligned probability curves, and in most cases completely absent. This gives us confidence that at least within the present paradigm, eye movement information does in fact present an adequate characterization of task selection dynamics.

We also tested the effect of response congruency (i.e., whether or not stimuli afforded the same or different responses across the two tasks) as a reflection of response selection demands on eye movement curves. Recall that we had found that generally a task switch decreased the tendency to move the eyes toward the target and increased the tendency to move toward the distractor (i.e., see Figure 6). Interestingly, response incongruency produced exactly the opposite pattern: There was a highly significantly greater tendency for eye movements to move toward the target on incongruent than on congruent trials, indicating that switch effects on eye movements represent more than just response-selection effects. Furthermore, this result also suggests that "early" attentional selectivity is modulated through conflict that arises in the "late" response-selection stage (Kuhns & Mayr, 2011). More generally, these results

---

[6] Another recent model by Altmann and Gray (2008) characterizes the memory processes that are involved in task switching, but because it contains no attentional parameters, it cannot be easily applied to the kind of data we present here.

suggest that eye movements inform not only about attentional selection, but they can provide indirect information about processes beyond the attentional selection stage.

Extensions

The analysis of the across-trial coupling patterns can help better understand control phenomena beyond the task-switching context. One currently prominent question in the literature is how, exactly, control is elicited. An important idea is that it is mainly the experience of conflict that triggers control attempts (e.g., Botvinick et al., 2001). Our results contain some relevant evidence. We found that trial $n-1$ eye movements toward the distractor led to an increase of target-directed eye movements on trial n (see Figure 7a). Importantly, this effect was not affected by switch transitions, suggesting that recruitment of control is general (Freitas et al., 2007) rather than task specific (Goschke, 2000; Monsell, Sumner, & Waters, 2003). As mentioned in the previous section, we also found evidence that experience of response conflict influences attentional selection within the same trial, which suggests a much more immediate regulatory process (see also Scherbaum, Fischer, Dshemuchadse, & Goschke, 2011) than the cross-trial modulation proposed in the original conflict adaptation model (Botvinick et al., 2001). Such results suggest that the present approach should be useful for future, fine-grained examinations of conflict adaptation processes.

Finally, in the context of the analysis of neuroimaging data, researchers have become very adept at analyzing within-subject temporal dynamics and dependencies. Initially, such analyses were driven by methodological necessity—to detect systematic effects, lagged dependences in the signal had to be modeled in an individual-specific manner. Increasingly, however, the within-individual dynamic coupling of activity patterns itself becomes an important source of information, in particular with respect to the functional connectivity of brain networks (Stephan et al., 2010). On the psychological/behavioral side, similar developments have been slower to come to the forefront. However, we show here that, with sufficient trial-by-trial data quality and adequate statistical methods, theoretically important information about behavioral level "functional connectivity" can be obtained. This approach should be useful for complementary neural level and behavioral analyses (Kievit, Romeijn, Waldorp, Wicherts, Scholte, & Borsboom, 2011). However, even just on the psychological/behavioral level, such an approach is appropriate whenever we are dealing with theories that describe a dynamic system and assuming sufficient within-subject data density is available. Beyond the area of cognition, domains where related developments are already underway include personality (Borsboom, Mellenberg, Van Heerden, 2003), emotion regulation (Thompson et al., in press), or psychiatric symptomatology (Cramer, Waldorp, van der Maas, & Borsboom, 2010).

Conclusion

We analyzed eye movements to assess how interference/control dynamics develop within and across trials in a task-switching situation. Within-trial dynamics revealed a discrete delay in the average onset of task-congruent attention on switch trials as long as the time to prepare was short and switch frequency was low to moderate. We interpret this pattern in terms of "special processes" that can be executed proactively and that are mandatory on switch trials, but are optional on no-switch trials. Likely, these involve retrieving task-relevant information into working memory (Mayr & Kliegl, 2003), setting attention on the now relevant dimension (e.g., Meiran, 2000), and possibly also disengaging from the previous task (e.g., Mayr & Keele, 2000). Also between-trial coupling dynamics were (with some qualifications) consistent with a model that assumes switch-specific, but task-unspecific, higher order configuration processes. To our knowledge, this is the first systematic assessment of how attentional settings influence one another across time and within individual subjects. We believe this provides a promising new way of characterizing the dynamics of control, one that complements more traditional, "static" approaches.

References

Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review, 115,* 602–639. doi:10.1037/0033-295X.115.3.602

Awh, E., Armstrong, K. M., & Moore, T. (2006). Visual and oculomotor selection: Links, causes and implications for spatial attention. *Trends in Cognitive Sciences, 10,* 124–130. doi:10.1016/j.tics.2006.01.001

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. doi:10.1016/j.jml.2007.12.005

Bates, D., Maechler, M., & Dai, B. (2008). Linear mixed-effects models using S4 classes. R package version 0.999375–27 [Software]. Vienna, Austria: R Foundation for Statistical Computing.

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review, 110,* 203–219.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652. doi:10.1037/0033-295X.108.3.624

Botvinick, M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to routine sequential action and its pathologies. *Psychological Review, 111,* 395–429. doi:10.1037/0033-295X.111.2.395

Brown, J. W., Reynolds, J. R., & Braver, T. S. (2007). A computational model of fractioned conflict-control mechanisms in task-switching. *Cognitive Psychology, 55,* 37–85. doi:10.1016/j.cogpsych.2006.09.005

Bryck, R., & Mayr, U. (2008). Task selection cost asymmetry without task switching. *Psychonomic Bulletin & Review, 15,* 128–134. doi:10.3758/PBR.15.1.128

Bunge, S. A., & Zelazo, P. D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science, 15,* 118–121. doi:10.1111/j.0963-7214.2006.00419.x

Cooper, R. F., & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review, 113,* 887–916. doi:10.1037/0033-295X.113.4.887

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers, 34,* 613–617. doi:10.3758/BF03195489

Cramer, A. O. J., Waldorp, L. J., van der Maas, H., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences, 33,* 137–193. doi:10.1017/S0140525X09991567

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

Dodd, M. D., Van der Stigchel, S., & Hollingworth, A. (2009). Novelty is

not always the best policy: Inhibition of return and facilitation of return as a function of visual task. *Psychological Science, 20,* 333–339. doi:10.1111/j.1467-9280.2009.02294.x

Findlay, J. M. (1997). Saccade target selection during visual search. *Vision Research, 37,* 617–631. doi:10.1016/S0042-6989(96)00218-0

Folk, C. L., & Remington, R. (2006). Top-down modulation of preattentive processing: Testing the recovery account of contingent capture. *Visual Cognition, 14,* 445–465. doi:10.1080/13506280500193545

Freitas, A. L., Bahar, M., Yang, S., & Banai, R. (2007). Contextual adjustments in cognitive control across tasks. *Psychological Science, 18,* 1040–1043. doi:10.1111/j.1467-9280.2007.02022.x

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multi-level/hierarchical models.* Cambridge, MA: Cambridge University Press.

Gilbert, S. J. (2005). Does task-set reconfiguration create cognitive slack? *Journal of Experimental Psychology: Human Perception and Performance, 31,* 92–100. doi:10.1037/0096-1523.31.1.92

Gilbert, S., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology, 44,* 297–337. doi:10.1006/cogp.2001.0770

Goschke, T. (2000). Involuntary persistence and intentional reconfiguration in task-set switching. In S. Monsell & J. Driver (Eds.), *Attention and Performance XVIII: Control of cognitive processes* (pp. 331–355). Cambridge, MA: MIT Press.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin, 136,* 849–874. doi:10.1037/a0019842

Kievit, R. A., Romeijn, J. W., Waldorp, L. J., Wicherts, J. M., Scholte, H. S., & Borsboom, D. (2011). Mind the gap: A psychometric approach to the reduction problem. *Psychological Inquiry, 22,* 67–87. doi:10.1080/1047840X.2011.550181

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition, 18,* 655–681. doi:10.1080/13506280902986058

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of inhibition in task switching: A review. *Psychonomic Bulletin & Review, 17,* 1–14. doi:10.3758/PBR.17.1.1

Krampe, R. T., Kliegl, R., Mayr, U., Engberg, R., & Vorberg, D. (2000). On the fast and slow of bimanual rhythm production: Parallel versus integrated timing. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 206–233. doi:10.1037/0096-1523.26.1.206

Kuhns, D., & Mayr, U. (2011). *Using eye movements to assess conflict-regulation dynamics.* Poster presented at the 52nd meeting of the Psychonomic Society, Seattle, WA.

Logan, G. D. (2003). Executive control of thought and action: In search of the wild homunculus. *Current Directions in Psychological Science, 12,* 45–48. doi:10.1111/1467-8721.01223

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. doi:10.1037/0096-1523.29.3.575

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108,* 393–434. doi:10.1037/0033-295X.108.2.393

Mayr, U. (2006). What matters in the cued task-switching paradigm: Tasks or cues? *Psychonomic Bulletin & Review, 13,* 794–799. doi:10.3758/BF03193999

Mayr, U. (2009). Sticky plans: Inhibition and binding during serial task control. *Cognitive Psychology, 59,* 123–153. doi:10.1016/j.cogpsych.2009.02.004

Mayr, U., & Bryck, R. L. (2005). Sticky rules: Integration between abstract rules and specific actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 337–350.

Mayr, U., & Keele, S. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129,* 4–26.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1124–1140. doi:10.1037/0278-7393.26.5.1124

Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 362–372. doi:10.1037/0278-7393.29.3.362

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1423–1442. doi:10.1037/0278-7393.22.6.1423

Meiran, N. (2000). Reconfiguration of stimulus task-sets and response task-sets during task-switching. In S. Monsell & J. Driver (Eds.), *Attention and Performance XVIII: Control of cognitive processes* (pp. 377–399). Cambridge, MA: MIT Press.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7,* 134–140. doi:10.1016/S1364-6613(03)00028-7

Monsell, S., & Mizon, G. A. (2006). Can the task-cueing paradigm measure an "endogenous" task-set reconfiguration process? *Journal of Experimental Psychology: Human Perception and Performance, 32,* 493–516. doi:10.1037/0096-1523.32.3.493

Monsell, S., Sumner, P., & Waters, H. (2003). Task-set reconfiguration with predictable and unpredictable task switches. *Memory and Cognition, 31,* 327–342. doi:10.3758/BF03194391

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science, 314,* 91–94. doi:10.1126/science.1127242

R Development Core Team. (2008). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. ISBN 3–900051-07–0. Retrieved from http://www.R-project.org

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

Rougier, N. P., & O'Reilly, R. C. (2002). Learning representations in a gated prefrontal cortex model of dynamic task switching. *Cognitive Science, 26,* 503–520. doi:10.1207/s15516709cog2604_4

Scherbaum, S., Fischer, R., Dshemuchadse, M., & Goschke, T. (2011). The dynamics of cognitive control: Evidence for within-trial conflict adaptation from frequency-tagged EEG. *Psychophysiology, 5,* 591–600.

Schneider, D. W., & Logan, G. D. (2005). Modeling task switching without switching tasks: A short-term priming account of explicitly cued performance. *Journal of Experimental Psychology: General, 134,* 343–367.

Schuch, S., & Koch, I. (2003). The role of response selection for inhibition of task sets in task shifting. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 92–105. doi:10.1037/0096-1523.29.1.92

Sporns, O., Chialvo, D., Kaiser, M., & Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *Trends in Cognitive Sciences, 8,* 418–425. doi:10.1016/j.tics.2004.07.008

Stephan, K. E., Penny, W. D., Moran, R. J., den Ouden, H. E. M., Daunizeau, J., & Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *NeuroImage, 49,* 3099–3109. doi:10.1016/j.neuroimage.2009.11.015

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica, 135,* 77–99. doi:10.1016/j.actpsy.2010.02.006

Thompson, R. J., Mata, J., Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Gotlib, I. H. (in press). Concurrent and prospective relations between attention to emotion and affect intensity: An experience sampling study. *Emotion.*

Van der Stigchel, S., Meeter, M., & Theeuwes, J. (2006). Eye movement trajectories and what they tell us. *Neuroscience & Biobehavioral Reviews, 30,* 666–679. doi:10.1016/j.neubiorev.2005.12.001

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference. *Psychological Bulletin, 136,* 601–626. doi:10.1037/a0019791

van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 746–759. doi:10.1037/0096-1523.30.4.749

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Back-Madruga, C., McPherson, S., . . . Miller, B. L. (2004). Relational integration and executive function in Alzheimer's disease. *Neuropsychology, 18,* 296–305. doi:10.1037/0894-4105.18.2.296

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. doi:10.1016/S0010-0285(02)00520-0

West, R., & Travers, S. (2008). Tracking the temporal dynamics of updating cognitive control: An examination of error processing. *Cerebral Cortex, 18,* 1112–1124. doi:10.1093/cercor/bhm142

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. doi:10.1037/0096-1523.29.2.455

Appendix

Trial-to-Trial Dynamics in the Gilbert and Shallice (2002) and the Configuration Models

We wanted to verify our intuitions about how autocorrelative patterns between successive control states play out in the Gilbert and Shallice model and in the configuration model (see Figures 2a and 2b). To this end, we adapted the original Gilbert and Shallice model that was used to simulate switching between Stroop color naming and word naming to our task situation. The basic architecture of the model involves two task-specific subnetworks for each of the two tasks (Task A and B, Figure 2a and 2b). All network units use sigmoid activation functions. Stimulus information is represented in terms of separate, task-specific features and fed to the input nodes. Weights for the links between input and output units are preset so that a stimulus feature for a particular task activates the corresponding output unit (e.g., S_{A1} to R_{A1}) and deactivates the alternative output unit (e.g., S_{A1} to R_{A2}). There are negative weights (i.e., inhibitory relationships) between the input for the same task (e.g., S_{A1} and S_{A2}) or between the output units for the same task (e.g., R_{A1} and R_{A2}). Corresponding response units for the two tasks (i.e., R_{A1} and R_{B1} or R_{A2} and R_{B2}) are linked in an excitatory manner, whereas noncorresponding response units (i.e., R_{A1} and R_{B2} or R_{A2} and R_{B1}) are linked in an inhibitory manner. This ensures that the model can produce response congruency effects where response times and error rates decrease when the stimulus affords the corresponding response options relative to when the stimulus affords noncorresponding response options.

For the carryover model, selection of appropriate tasks is achieved through "task demand" units that also use exponential activation functions and of which there is one for each of the two tasks. The task demand units receive discrete, on/off input depending on which task is currently relevant, and they are connected with each other through mutually inhibitory links. The task demand units provide excitatory activity to the corresponding response units and inhibitory activity to the noncorresponding response units. Appropriate output activity from the task demand units moves corresponding response unit activation levels into a region where they become sensitive to input unit activity. Thus, the

demand units function as a gate between input and output unit activity. As a mechanism for bottom-up-driven task set interference, task demand units also receive input from response units. Carryover of task demand activity is achieved in two ways: (a) by preserving some of the task demand activity between trial $n-1$ and trial n; (b) through single-trial Hebbian learning between response input unit and task demand units, which is reset after each trial and that leads to a tendency to reactivate the previous trial task demand unit.

The model we used differs from the original Gilbert and Shallice model in a few aspects. First, the original model used three instead of only two response options per task (see also Gilbert, 2005). Second, whereas for our model both tasks were equally "strong," the original model implemented an asymmetric dominance relationship between their two tasks (established via stronger stimulus input for the dominant vs. the nondominant task, and compensatory stronger control input for the task demand units of the nondominant task than the dominant task), which was not the case for our model. Third, instead of the alternate runs paradigm simulated by Gilbert and Shallice, tasks were randomly selected ($p = .50$ switch rate; variations in switch probability did not produce different results, which is not surprising, given that the model has no mechanisms for adapting to such manipulations). We confirmed that these changes to the original model did not alter the qualitative pattern of mean effects reported by Gilbert and Shallice (2002).

The configuration model required the following additional changes. The task-specific carryover parameter, the Hebbian learning between stimulus and task demand units, and the weights governing mutual inhibition among task demand units were set to zero. Instead, we added a "high-level control input" that provided a positive top-down bias to the currently relevant and an equal sized negative bias to the currently irrelevant task demand unit. The size of the high-level control input (C) for trial n was given by equation:

$$C_n = 6 - (6 - C_{n-1}) \times .75 + noise,$$

where the initial value for C was 6, and *noise* was a rectangular random distribution between $-.5$ and $+.5$, ensuring an autocorrelation between successive trial control states.

To characterize the trial-to-trial coupling pattern of the task-specific carryover and the configuration model, we ran the simulation for 10,000 trials with short CSI. We present in the left panels of Figure 3 the correlations between the activity for successive trial's target and distractor task demand units after 20 cycles within a trial and also with simulated RTs. The right panels of Figure 3 contain the corresponding values for the configuration model. The top and the bottom panels use alternative specifications for analyzing trial-to-trial relationships (see text for details). We chose the

activity level after 20 cycles to represent the attentional setting early in the trial; however, results did not change in a qualitative manner when using activation patterns after different numbers of cycles. We used simple correlations to portray the pattern of trial-to-trial coupling because in the model, the different trial $n-1$ predictors are highly intercorrelated (e.g., there is no reliable relationship between trial $n-1$ predictors and trial n RTs after controlling for trial n predictors).