

Drifting from slow to “D’oh!” Working memory capacity and mind wandering predict extreme reaction times and executive-control errors.

By: Jennifer C. McVay and Michael J. Kane

[McVay, J.C.](#), & [Kane, M.J.](#) (2012). Drifting from slow to “D’oh!” Working memory capacity and mind wandering predict extreme reaction times and executive-control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 525-549. [DOI: 10.1037/a0025896](#)

Made available courtesy of the American Psychological Association. This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

Abstract:

A combined experimental, individual-differences, and thought-sampling study tested the predictions of executive attention (e.g., Engle & Kane, 2004) and coordinative binding (e.g., Oberauer, Süß, Wilhelm, & Sander, 2007) theories of working memory capacity (WMC). We assessed 288 subjects' WMC and their performance and mind-wandering rates during a sustained-attention task; subjects completed either a go/no-go version requiring executive control over habit or a vigilance version that did not. We further combined the data with those from McVay and Kane (2009) to (1) gauge the contributions of WMC and attentional lapses to the worst performance rule and the tail, or τ parameter, of reaction time (RT) distributions; (2) assess which parameters from a quantitative evidence-accumulation RT model were predicted by WMC and mind-wandering reports; and (3) consider intrasubject RT patterns—particularly, speeding—as potential objective markers of mind wandering. We found that WMC predicted action and thought control in only some conditions, that attentional lapses (indicated by task-unrelated-thought reports and drift-rate variability in evidence accumulation) contributed to τ , performance accuracy, and WMC's association with them and that mind-wandering experiences were not predicted by trial-to-trial RT changes, and so they cannot always be inferred from objective performance measures.

Keywords: executive control | experimental psychology | psychology | individual differences | mind wandering | reaction time | working memory | reaction times | cognitive processes | executive function

Article:

Acknowledgement: Portions of this work were supported by a National Institutes of Health Ruth L. Kirschstein National Research Service Award (NRSA) Individual Predoctoral Fellowship granted to Jennifer C. McVay (Award Number F31MH081344) from the National Institute of Mental Health (NIMH). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIMH. We are grateful to Cody Burns, DeLaura

Jansen, Meaghan Kerns, Dara Rogers, Lindsay Squires, Alanna Stockford, Meaghan X, and Hanna Zombek for their assistance in data collection.

Correspondence concerning this article should be addressed to: Jennifer C. McVay, Department of Psychology, University of North Carolina at Greensboro, P.O. Box 26170, Greensboro, NC 27402-6170 Electronic Mail may be sent to: jennifercmcvcay@gmail.com.

Correspondence concerning this article should be addressed to: Michael J. Kane, Department of Psychology, University of North Carolina at Greensboro, P.O. Box 26170, Greensboro, NC 27402-6170 Electronic Mail may be sent to: mjokane@uncg.edu.

People tend to make mistakes when they think too much (e.g., Beilock & Carr, 2001) or too little (e.g., Reason, 1990) about ongoing, routine activities. The present study explores whether executive control over thought content—and over mind wandering, in particular—contributes to individual differences in working memory capacity (WMC) and their cognitive and behavioral consequences. Attentional theories of WMC argue that domain-general, executive control capabilities contribute to performance on both WMC and higher order cognitive tasks, as well as to their shared variance (e.g., Braver, Gray, & Burgess, 2007; Hasher, Lustig, & Zacks, 2007; Hasher & Zacks, 1988; Kane, Conway, Hambrick, & Engle, 2007; Unsworth & Engle, 2007; Unsworth & Spillers, 2010). Some evidence for these views comes from studies showing that WMC measures predict not only complex cognitive skills, such as reasoning and reading (e.g., Daneman & Merikle, 1996; Kane, Hambrick, & Conway, 2005; Oberauer, Schulze, Wilhelm, & Süß, 2005), but also more simple attention functions, such as restraining habitual but contextually inappropriate responses (e.g., Hutchison, 2011; Kane & Engle, 2003; Long & Prat, 2002; Unsworth, Schrock, & Engle, 2004) or constraining conscious focus to target stimuli amid distractors (e.g., Colzato, Spapé, Pannebakker, & Hommel, 2007; Conway, Cowan, & Bunting, 2001; Fukuda & Vogel, 2009; Heitz & Engle, 2007; Poole & Kane, 2009; but see Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008; Keye, Wilhelm, Oberauer & van Ravenzwaaij, 2009).

Our view of executive control, like others' (e.g., Braver et al., 2007; Cohen & Servan-Schreiber, 1992; De Jong, 2001; Jacoby, Kelley, & McElree, 1999; Roberts & Pennington, 1996), is that self-regulation of thought and behavior is sometimes accomplished proactively, in advance of stimuli or contexts that provoke distraction, conflict, or other challenges. We have proposed that proactive control is accomplished by the active maintenance of goal representations (Engle & Kane, 2004; Kane, Conway, et al., 2007): If goals are not kept accessible, then strong distractors or habits may inappropriately capture ongoing cognition and performance, resulting in goal neglect errors (Duncan, 1995) and action slips. We also argue that goal maintenance, which varies with WMC, is fragile and can be disrupted by salient external stimuli or by task-unrelated thoughts (TUTs) that are mentally or environmentally cued (McVay & Kane, 2009, 2010, 2011). By this view, individual differences in the ability to maintain on-task thoughts may contribute to WMC's effects on attention-task performance. The online assessment of TUTs, then, provides a potentially powerful method to test our executive attention theory against those that do not hold

attention control to be a significant source of WMC variation or its covariation with other capabilities (e.g., Colom et al., 2008; Mogle, Lovett, Stawski, & Sliwinski, 2008; Oberauer, Süß, Wilhelm, & Sander, 2007).

WMC, Goal Neglect, and Executive Control

Tasks requiring subjects to occasionally withhold prepotent responses in favor of novel ones provided initial evidence that WMC predicted goal neglect. In the antisaccade task, for example, subjects with higher WMC scores better resisted the lure of a flashing visual cue in order to orient their attention away from the cue and toward the target, which always appears in the opposite direction. On these antisaccade trials, higher WMC is associated with greater target-identification accuracy and fewer erroneous, “joked,” saccades toward the cues (Kane, Bleckley, Conway, & Engle, 2001; Unsworth et al., 2004). Higher WMC subjects thus seem to keep goal representations more accessible than do lower WMC subjects, allowing those goals to better guide behavior in the moment.

In subsequent work with the Stroop task, we manipulated the extent to which the task context reinforced the color-naming goal. Our idea was that WMC should predict goal neglect, especially in situations that put a premium on proactively maintaining goal access (Kane & Engle, 2003; see also Marcovitch, Boseovski, Knapp, & Kane, 2010). We therefore presented mostly congruent trials (75%–80% of trials), in which words were presented in their matching colors (e.g., RED in red), along with explicit instructions to continue ignoring the words even if they often matched their color. This high-congruent context thus allowed subjects who failed to maintain the color-naming goal to nonetheless respond accurately on most trials based on a word-reading habit. Indeed, lower WMC subjects more often slipped into word reading than did higher WMC subjects, leading them to commit significantly more errors on infrequent incongruent trials (and to respond especially quickly on congruent trials, also suggestive of word reading). We found no such accuracy differences between WMC groups in low-congruency contexts, where most trials presented color–word conflict, thus reinforcing the ignore-the-word goal. We thus argued that, in the absence of strong external support, subjects with lower WMC will more often lose access to task goals and commit habit-based errors than will higher WMC subjects.

Kane and Engle (2003) further suggested that, even though high-congruency contexts yielded significant WMC differences in accuracy, evidence for goal neglect might also arise in subjects' longest reaction times (RTs; see also De Jong, 2001). That is, occasional long RTs might reflect momentary, incomplete failures of goal maintenance that are corrected just in time (perhaps in response to conflict-detection mechanisms; e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001). These arguments paralleled those previously made in the intelligence literature regarding individual differences in long RTs (i.e., the worst performance rule; Larson & Alderton, 1990). The worst performance rule describes the fact that the longest RTs that people sometimes commit in choice-RT tasks are more strongly correlated with their cognitive ability (e.g., fluid

intelligence) than are the shortest RTs that they are able to commit. That is, when individual subjects' RTs are ranked from shortest to longest and averaged into quantile bins, RT–intelligence correlations increase steadily with increasing RT quantiles (for a review, see Coyle, 2003). According to many researchers, fluctuations in working memory maintenance (Larson & Alderton, 1990; Larson & Saccuzzo, 1989) or in attention to the task (Jensen, 1992) explain the worst performance rule (but see Ratcliff, Schmiedek, & McKoon, 2008). Momentary lapses in task focus result in especially long RTs to affected trials, and people with lower intelligence suffer more of these lapses. This explanation is, of course, consistent with the executive attention theory of WMC, according to which high-WMC subjects have better goal-maintenance capabilities and therefore commit fewer long RTs due to lapses of attention than do low-WMC subjects.

Unsworth, Redick, Lakey, and Young (2010) tested whether the worst performance rule applied to WMC by examining vigilance-task RTs by ranked bins, as is typical in such studies. They also quantified worst performance by considering the tail of the positively skewed RT distribution. Ex-Gaussian models represent individual subjects' RT distributions as a convolution of a Gaussian distribution and an exponential distribution; such models have three parameters: the mean and standard deviation of the Gaussian component; μ (μ) and σ (σ); and the mean of the exponential, tail component, τ (τ). By the worst performance rule, Unsworth et al. expected WMC to predict the tail, or τ , more strongly than the leading edge, or μ , of the RT distribution. Indeed, τ is sensitive to experimental manipulations of, and age-related differences in, executive control (e.g., De Jong, Berendsen, & Cools, 1999; Tse et al., 2010; West, 2001); variation in the τ parameter may therefore reflect, at least in part, periodic lapses of attention to task goals. Unsworth et al. used latent-variable analyses to derive a WMC factor from multiple tasks and to test its relation to vigilance-task RT quintiles and ex-Gaussian parameters. WMC correlated more strongly with longer than with shorter RT quintiles and more strongly with τ than with μ . Unsworth et al. thus concluded that WMC-related variation in “worst performance,” or long RTs, reflected variation in susceptibility to attentional lapses or TUTs, consistent with executive attention theory (e.g., Kane, Conway, et al., 2007; Unsworth & Spillers, 2010).

An Alternative “Binding” View of WMC and Long RTs

Yet other interpretations are possible. For example, the coordinative binding theory attributes individual differences in WMC, as well as WMC's relation to intellectual ability, to a limited capacity for temporary, simultaneous bindings of distinct mental representations into coherent cognitive structures (e.g., Oberauer, 2005, 2009; Oberauer et al., 2007). Oberauer and colleagues have argued that the evidence for a strong association between WMC and executive control is not yet compelling, noting that individual differences in WMC variation only weakly predict task-set-switching costs, which arguably mark deficiencies in executive control (Oberauer et al., 2007; but for an alternative view, see Kane, Conway, et al., 2007). Moreover, coordinative binding theory explains the empirical associations between WMC and response-conflict tasks, such as antisaccade and Stroop, via the demands on stimulus–response (S-R) binding, rather than

on executive control (Oberauer, 2009; Wilhelm & Oberauer, 2006). In such “attention” tasks, the critical trials require highly incompatible S–R mappings (e.g., if a flash appears to the right, look left; if RED appears in blue, say “blue”). According to binding theory, lower WMC subjects' difficulties here result from deficiencies in temporarily binding the task-relevant stimuli onto arbitrary response rules and maintaining those novel bindings throughout the task. Thus, binding deficits—and not control failures—leads to problems in response selection and performance.

Based on this binding view, Schmiedek, Oberauer, Wilhelm, Süß, and Wittmann (2007) also presented an alternate account of the worst performance rule, whereby lapses of attention do not contribute to the relation between long RTs and cognitive ability (e.g., WMC). Rather, WMC-related differences in establishing and maintaining S–R bindings lead to differences in the “efficiency of information transmission between stimuli and responses” (p. 425), and generally poor efficiency produces occasionally long RTs (see also Martínez & Colom, 2009). Schmiedek et al. used a version of Ratcliff's diffusion model to assess information-processing efficiency (drift rate of the evidence-accumulation process; e.g., Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004) and then examined the association between the efficiency/drift parameter and τ . In brief, the diffusion model is a random-walk, evidence-accumulation model that quantitatively accounts for choice-task accuracy and RT data (including RT distributions), typically via seven or eight main parameters of interest that correspond to between- and within-subject processing variables 2; it thus models both group- and individual-level data.

Schmiedek et al. (2007) derived μ , σ , and τ latent factors from the ex-Gaussian parameter estimates across eight choice-RT tasks and used structural equation modeling to test relations of these factors to a WMC factor based on six tasks. Only the τ factor predicted unique variance in WMC ($\beta = -.90$). For diffusion modeling, however, given the limited number of trials per RT task, Schmiedek et al. used a reduced, “EZdiffusion” model (Wagenmakers, van der Maas, & Grasman, 2007) that estimates only three parameters: (1) the response criterion, or initial distance between the start point and decision boundaries; (2) the nondecision parameter, or time spent on nondecision processes (e.g., stimulus encoding, response execution); and (3) the drift rate, or mean rate at which evidence accumulates toward a boundary (for Schmiedek et al., 2007, drift rate reflected the general quality of the information processing). The authors proposed that τ is driven primarily by drift rate and that drift rate accounts for the WMC– τ correlation. Indeed, Schmiedek et al. found that the WMC– τ association was of similar magnitude to the WMC–drift rate association. Moreover, a subsequent simulation study took the EZdiffusion parameter values derived from the RT data, simulated new RTs based on only those parameters, and successfully reproduced the original WMC– τ correlation. It appears, then, that only three parameters were necessary to explain WMC-related variation in RT, none of which corresponded closely to attentional lapses. The evidence thus suggested that WMC may predict long RTs in simple tasks without appealing to any influences of attentional lapses or mind-wandering (TUT) experiences.

Schmiedek et al. (2007) acknowledged that they could not rule out some (potentially minor) contribution of attentional lapses to individual differences in τ , worst performance, and WMC variation therein. They also argued, however, that additional positive evidence for such an attentional contribution was needed because their EZdiffusion modeling results, which required no attentional-lapse parameter, provided a more parsimonious account. The main goal of the present study was to test for just that positive evidence for attentional lapses influencing long RTs and WMC's association with them.

Our logic was that attentional lapses can be measured, albeit imperfectly, via thought probes that ask subjects to report whether their immediately preceding thoughts were on- or off-task (for a review, see Smallwood & Schooler, 2006). By probing for TUTs, we could test whether variation in off-task thinking was at all associated with WMC, τ , or both, and whether subjects' mind-wandering rates might account for some of the shared variance between WMC and τ . If individual differences in τ reflect, in part, diffusion-model drift rate (or something like general processing efficiency), but not the effects of attentional lapses (Schmiedek et al., 2007), then τ estimates should be uncorrelated with subjects' mind-wandering rates. If, however, TUT rate correlated with τ , and furthermore, if TUT rate mediated the association between WMC and τ , then this would provide novel evidence for the attentional-lapse explanation of the worst performance rule and of WMC's prediction of long RTs.

As well, the Schmiedek et al. (2007) inferential argument against attentional lapses was based on a reduced diffusion model that lacked a parameter corresponding to lapses. We suggest, along with Schmiedek et al., that attentional lapses might be captured by the diffusion model's parameter η , or across-trial variation in drift rate. That is, subjects who more often flow between on- and off-task thought during a task should show more variation across trials in information accumulation, or drift rate. Because the EZdiffusion model has no drift-rate variability parameter, absorbing any effects of drift variability into drift rate (Wagenmakers et al., 2007), the Schmiedek et al. argument from parsimony is weaker than it otherwise might be. Here, then, we modeled our subjects' RT data using a quantitative evidence-accumulation model of choice that included parameters for both drift rate and its variability (the linear-ballistic accumulator [LBA] model; Brown & Heathcote, 2008) and tested whether individual differences in either parameter captured any WMC or TUT effects on RT task performance.

It is worth noting, however, that we do not claim that the τ parameter from ex-Gaussian models can be identified with any particular, singular cognitive process or ability, such as vulnerability to attentional lapses. As Matzke and Wagenmakers (2009) established, researchers have identified a remarkable breadth of experimental manipulations that seem, within the confines of any one investigation, to selectively affect τ . Theorists thus have proposed a variety of different cognitive processes that supposedly characterize, or give rise to, τ . Indeed, Matzke and Wagenmakers further demonstrated, via simulations and empirical work, that τ is sensitive to at least two different diffusion-model parameters, corresponding to the theoretical processes of evidence accumulation and criterion setting. Viewed collectively, then, the choice-RT literature

indicates that τ (or worst performance) is not caused by a single process or mechanism. This really should not be surprising, given that slow responses must mean different things across different tasks that make unique cognitive demands, across different subjects who vary in abilities and motivations, and across varied contexts that afford a wealth of strategic approaches to tasks. Our purpose in this study was thus to test (contra Schmiedek et al., 2007) whether individual differences in vulnerability to attentional lapses play any contributing role—not the only role—in producing normal variation in τ , or very long RTs, within a particular long-duration, executive control task that appears to elicit significant mind wandering (McVay & Kane, 2009).

WMC, Mind Wandering, and Task Performance

In line with executive attention theory (e.g., Engle & Kane, 2004; Kane, Conway, et al., 2007), WMC variation predicts the propensity for mind wandering during cognitively demanding tasks and activities in daily life (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Gathercole et al., 2008; Kane, Brown et al., 2007). But does this WMC-related variability in TUTs contribute to empirical associations between WMC and task performance, including long RTs? As a preliminary test of executive attention theory versus coordinative binding theory, McVay and Kane (2009) administered WMC tasks and a go/no-go task (the sustained attention to response task [SART]; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997) that featured periodic thought probes to assess subjects' thoughts (as on- vs. off-task) in the moments before critical no-go trials. Binding theory claims no link between WMC and attentional lapses, nor a contribution for mind wandering to WMC-related variation in executive task performance. In fact, McVay and Kane found that subjects generally made more errors on trials on which they reported off-task versus on-task thoughts, and individual differences in TUT rate predicted overall SART performance and WMC. Furthermore, in regression analyses, TUT rate partially mediated the relation between WMC and SART performance. All of these correlational results suggest a role for thought control in executive control more broadly, and they support executive attention theory more specifically. McVay and Kane argued that it was not clear how limitations in capacity for temporary bindings (Oberauer et al., 2007) could account for a greater incidence of mind wandering in low versus high WMC individuals or, most critically, for these TUT rates contributing to WMC correlations with task performance.

At the same time, the McVay and Kane (2009) regression results indicated that TUT rate only partially mediated WMC's effects on performance. That is, WMC accounted for significant SART variance beyond that shared with mind wandering, and so attentional lapses could not fully explain WMC variation or covariation. We thus argued, based on the dual-component executive attention theory (Engle & Kane, 2004; Kane, Conway, et al., 2007; see also Braver et al., 2007), that the additional, TUT-independent variance was attributable to the more reactive competition-resolution component of executive control rather than the more proactive goal-maintenance component. The SART's frequent go trials and rare no-go trials build up a habitual tendency to respond rather than withhold responding. This creates additional interference on the

no-go trials such that, even when the goal of the task is proactively maintained, subjects still experience in-the-moment response conflict and sometimes produce incorrect responses. Thus, the SART, like the Stroop task (Kane & Engle, 2003), may be sensitive to WMC variation due to the premium it puts on both goal maintenance and competition resolution. According to McVay and Kane, then, individual differences in competition resolution explain the WMC-related variance in SART performance that is unrelated to TUT rate:

We therefore suggest that WMC's TUT-independent prediction of SART performance is largely due to its relation to competition resolution. If so, two predictions follow: (a) A SART that induces weaker prepotencies to overcome should correlate less strongly with WMC (due to a minimization of competition-resolution variance) and (b) SART variance that is predicted by WMC should be more fully mediated by TUT rate, as subjects must maintain goal activation that is not externally reinforced. (McVay & Kane, 2009, p. 203)

Thus, a primary goal of the current study was to assess mind wandering and its consequences in a version of the SART that made less demand on competition-resolution processes. Here we administered two versions of the SART in a between-subjects design: The standard SART, with infrequent no-go targets, and a vigilance SART, with infrequent go targets. The vigilance SART, then, like classic vigilance tasks (see Davies & Parasuraman, 1982), required subjects to withhold responses to most trials and wait to respond overtly to rare targets. By simply reversing the proportions of go and no/go trials, we removed the habit-inducing go response from the task and made it primarily dependent on goal maintenance rather than competition resolution. We predicted, therefore, that without additional demands for competition resolution, TUT variation (and attendant disruptions to goal maintenance) would fully mediate the association between WMC and vigilance-SART performance. Note also that current instantiations of binding theory (e.g., Oberauer et al., 2007) similarly predict WMC to correlate with standard SART performance beyond any influence of TUT rate, due to WMC-related variation in S–R binding. However, the S–R rules for the standard and vigilance SART were identical (press a key in response to animal names only), and so binding theory would seem to predict similar WMC associations to performance of both SART versions and little influence of TUT rate in either. 3

Objective Markers of Mind Wandering in RTs?

Mind wandering, as the subjective experience that accompanies attentional lapses, is typically measured subjectively: Subjects describe or classify their immediately preceding thoughts or experiences at periodic probes that briefly interrupt the ongoing primary task (Smallwood & Schooler, 2006). Despite their introspective nature, subjects' immediate thought reports appear to be reasonably valid. Such validity is not unexpected, given that these verbal reports describe contents of experiences rather than interpretations or explanations of those experiences (Nisbett & Wilson, 1977; Wilson, 2002). Across multiple laboratories, probed TUT reports vary consistently with (1) experimental treatments and task demands (e.g., Antrobus, Singer, & Greenberg, 1966; Forster & Lavie, 2009; Giambra, 1989, 1995; Grodsky & Giambra, 1990–

1991; McKiernan, D'Angelo, Kaufman, & Binder, 2006; Stuyven & Van der Gooten, 1995; Teasdale et al., 1995; Teasdale, Proctor, Lloyd, & Baddeley, 1993); (2) practice and time on task (e.g., Antrobus, Coleman, & Singer, 1967; Antrobus et al., 1966; McVay & Kane, 2009; Perry & Laurie, 1992; Smallwood et al., 2004; Smallwood, Obonsawin, & Reid, 2003, Smallwood, Riby, Heim, & Davies, 2006; Teasdale et al., 1995); (3) subjective and objective measures of task performance (McVay & Kane, 2009; McVay, Kane, & Kwapil, 2009; Schooler, Reichle, & Halpern, 2004; Smallwood, McSpadden, Luus, & Schooler, 2008; Smallwood, McSpadden, & Schooler, 2007; Smallwood, McSpadden, & Schooler, 2008); (4) neuroimaging signatures (e.g., Christoff, Gordon, Smallwood, Smith, & Schooler, 2009; Mason et al., 2007; McGuire, Paulusu, Frackowiak, & Frith, 1996; McKiernan et al., 2006; Smallwood, Beach, Schooler, & Handy, 2008); and (5) individual-differences variables (e.g., Giambra, 1989, 1993; Kane, Brown, et al., 2007; McVay & Kane, 2009; Shaw & Giambra, 1993; Smallwood et al., 2002–2003; Smallwood, O'Connor, Sudbery, & Obonsawin, 2007). Nonetheless, the inherent subjectivity of thought reports has led investigators to seek objective behavioral markers of attentional lapses and TUT experiences.

Some authors have suggested that neglect-type performance errors may serve as objective measures of TUTs, at least in some task contexts (Cheyne, Solman, Carriere, & Smilek, 2009; Smallwood, Beach, et al., 2008; Smallwood, Fitzgerald, Miles, & Phillips, 2009; Smallwood, McSpadden, et al., 2008; but see Smallwood et al., 2006). Empirical evidence indicates, however, that TUTs and errors are not always interchangeable. Although they may elicit somewhat similar event-related potentials (Smallwood, Beach, et al., 2008), goal-neglect errors and TUT reports can vary independently. Smallwood, McSpadden, et al. (2008) found that faster stimulus presentation rates reduced TUTs but not errors. As well, McVay and Kane (2009) found that errors increased when subjects reported mind wandering versus on-task thinking, but mean error rate during TUTs did not approach 100% (it was 68%); moreover, the error rate during reports of on-task thinking was a sizeable 34%. Theoretically, of course, SART errors may arise not only from goal neglect but also from insufficient resolution of response competition, from momentary speed–accuracy shifts, or from myriad other sources (see also Helton et al., 2005; Helton, Weil, Middlemiss, & Sawers, 2010). Errors and mind wandering, therefore, are not likely to be isomorphic in all contexts, and so they should not be treated as interchangeable without independent evidence.

Might transient changes in RTs provide a more consistent objective signal of lapses? Robertson et al. (1997) observed a speedup of RTs in the trials immediately preceding a target-trial error in the SART. Although they suggested that these shorter RTs predict errors because they reflect mindless responding, Robertson et al. did not assess TUTs via thought reports. McVay and Kane (2009) similarly found that shorter RTs preceded both errors (vs. accurately withheld target responses) and TUTs (vs. on-task thought reports; see also Smallwood et al., 2004; Smallwood, McSpadden, et al., 2008). However, as yet another indication of the divergence between objective and subjective measures of attentional lapses, the magnitude of the RT difference on

trials preceding errors versus accurate responses ($M = 73$ ms) was much greater than in trials preceding TUTs versus on-task thoughts ($M = 11$ ms), suggesting that there can be more to these short pre-error RTs than just mind wandering. Indeed, Jackson and Balota (2011) have also found that older and younger adults differ dramatically in self-reported TUT rates during the SART (with older adults reporting much less mind wandering), and yet both younger and older adults speed up similarly in the trials that occur just before no-go errors. Thus, in the same way that errors may have multiple causes, an RT speedup in the SART may reflect not only TUTs but also (or instead) a buildup of motor habit or a speed–accuracy tradeoff.

Smallwood, McSpadden, et al. (2008) undertook a more sophisticated investigation of SART RT sequences, searching for consistent time-series patterns and testing whether any predicted imminent errors or TUTs. Their SART manipulated presentation rate and block length, and each block terminated in either a thought probe or a no-go target. Using principal components analysis (PCA) on RTs from the 12 nontargets that preceded block endings, Smallwood, McSpadden, et al. extracted three components onto which all RT sequences loaded. 4 The components represented, in order of variance accounted for, (1) general RT, or the extent to which each run of 12 was faster or slower than average; (2) linear RT change (slowing or speeding) just prior to the probe or target that terminated the block; and (3) quadratic RT change (from slower to faster to slower, or vice versa) just prior to probe or target. In fact, the outcomes of interest were modestly predicted by some of these components. Component 2 had significantly higher scores on trials preceding an error compared with a baseline (i.e., all blocks terminating in thought probes). The runs preceding on-task thought reports, in contrast, had lower Component 2 scores than baseline (i.e., all blocks terminating in a target). Finally, blocks ending in “zone-out” reports (i.e., TUTs without one’s prior awareness; Schooler, 2002) had lower Component 1 scores than did those ending in on-task reports.

Smallwood, McSpadden, et al. (2008) claimed that the discernible RT patterns prior to on- versus off-task thought reports may objectively mark TUTs. Unfortunately, in the key analyses of Components 1 and 2, the differences in component scores involved a potentially contaminated baseline. That is, the baseline to which they compared Component 2 scores for on-task thought reports included both correct and error trials, which had elevated Component 2 scores in an analysis comparing errors to “baselines” that included both on- and off-task thought reports; this circularity would need to be broken in order to unambiguously interpret their findings. As well, Smallwood, McSpadden, et al. intended to measure within-subject, in-the-moment RT patterns that predicted performance and thoughts, but they failed to standardize RTs within-subject, and so between-subjects differences in RTs may have had undue influence on the principal components (see, e.g., Klinger & Cox, 1987–1988). That is, rather than indicating that TUTs are more likely to occur on occasions when a subject speeds up before a probe, their findings may have indicated that TUTs are more likely to occur for subjects who tend to speed up. Thus, as a secondary goal of this study, we attempted to replicate the RT patterns from Smallwood, McSpadden, et al. with a stronger analytic approach.

Summary of Experimental Aims

In a single, large- N study we examined the executive attention (Engle & Kane, 2004; Kane, Conway, et al., 2007) and coordinative binding (Oberauer et al., 2007; Wilhelm & Oberauer, 2006) theories of WMC in two ways. In both, we relied on probed thought reports to measure attentional lapses as TUT experiences. First, we tested the extent to which individual differences in TUT rate mediated the WMC–SART association, when SART performance relied heavily on both goal-maintenance and competition-resolution processes (in the standard, go/no-go SART) versus when it relied little on competition resolution (in the vigilance SART). Executive attention theory predicted that TUT rate would partially mediate the WMC–performance relation when both goal maintenance and competition resolution were needed (replicating McVay & Kane, 2009) but that TUT rate would more fully mediate the WMC–performance relation when only goal maintenance was needed. Second, we asked whether, by measuring attentional lapses via thought probes and modeling SART RTs with an evidence-accumulation model that assessed drift-rate variability, we would find evidence that the worst performance rule in RTs was at all influenced by attentional lapses, as claimed by executive attention theory (see Unsworth et al., 2010). The coordinative binding theory, in contrast, claims that worst performance, and WMC's association with it, reflects general information-processing efficiency (indicted by drift rate) resulting from S–R binding effectiveness rather than attentional lapses (Schmiedek et al., 2007). Finally, we used RTs as a predictor rather than an outcome of interest by investigating whether latency time-series data, particularly series preceding subjects' performance errors or TUT reports, might provide converging, objective evidence for attentional lapses. Here we specifically tested whether the Smallwood, McSpadden, et al. (2008) findings, which appear so promising in this regard, would replicate in a substantially larger data set and with an improved analytic approach.

WMC, Mind Wandering, and SART Demands

Method

Subjects

Two hundred eighty-eight undergraduates (18–35 years of age) at the University of North Carolina at Greensboro (UNCG) participated in two sessions, one for testing WMC and one for completing the SART. We dropped data from four subjects who didn't follow SART instructions (two each in the standard and vigilance tasks). Due to experimenter error, three subjects did not complete WMC testing (one from the standard and two from the vigilance groups); we included their SART data for all non-WMC-related analyses.

General procedure

We tested subjects in groups of one to six. They completed the WMC and SART sessions within the same semester ($M = 24$ days between sessions, $SD = 26$, range = 1–95). During the first, WMC screening session, subjects completed the operation span (OSPAN), symmetry span (SSPAN), and reading span (RSPAN) tasks in that order. During the second session, subjects completed the SART.

WMC screening

In 90-min sessions, subjects completed three automated complex-span tasks: OSPAN, RSPAN, and SSPAN. The tasks required subjects to maintain access to memory items while completing an unrelated processing task; the processing task had an individualized response-latency deadline ($M + 2.5$ SDs), calculated during 15 processing-task-only practice items (Unsworth, Heitz, Schrock, & Engle, 2005; Unsworth, Redick, Heitz, Broadway, & Engle, 2009). Each trial of the processing task was followed, 200 ms later, by a memory item. In OSPAN, subjects verified compound equations while remembering letters (from a pool of 12), each presented for 250 ms following every equation. For example, following a screen with “ $(2*2) + 1 = 5$,” the subject would click a response (TRUE or FALSE) with a computer mouse and then see the to-be-remembered letter, “F,” onscreen. RSPAN similarly required subjects to remember letters but while verifying the meaningfulness of sentences (e.g., “I like to run in the sky”). In SSPAN, subjects verified vertical symmetry of black-and-white matrix patterns while remembering the locations of red squares on a 4×4 grid. Each red square was presented for 650 ms after the symmetry problem.

The WMC tasks presented each set length (3–7 in OSPAN and RSPAN; 2–5 in SSPAN) three times, randomly ordered for each subject. At the end of each set, subjects recalled the memory items in serial order. For OSPAN and RSPAN, subjects used a mouse to click on that set's letters in order, presented among the full pool of 12 letters. For SSPAN, subjects used a mouse to click the previously occupied squares in order, within an empty 4×4 grid.

The span score for each task was the sum of items recalled in serial position (of 75 in OSPAN and RSPAN and 42 in SSPAN; Conway et al., 2005). We converted span scores to z scores (based on the Ms and SDs from our UNCG database of over 2,000 undergraduates) and averaged them into a WMC composite. Scores correlated as $r = .64$ (RSPAN \times OSPAN), $r = .52$ (OSPAN \times SSPAN), and $r = .47$ (SSPAN \times RSPAN). The WMC variable was normally distributed (skew = -0.67 ; kurtosis = 0.14).

SART session

Materials and design

The design was a 2×2 mixed-model factorial, with SART type (standard vs. vigilance) manipulated between subjects and trial type (target, nontarget) within subject. In the standard SART ($N = 142$), subjects responded to frequent nontarget words and withheld responses to

infrequent (11%) target words; thus, as is standard for the SART, it was a go/no-go task that elicited a go prepotency. In the vigilance SART (N = 142), in contrast, subjects responded to only the infrequent (11%) targets; it was therefore a prototypical vigilance task without any go prepotency. For both SARTs, the words were from two different categories (foods and animals; e.g., animals as targets and foods as nontargets). Stimuli appeared in black against a white background, in 18-point Courier-New font, via CRT or LCD monitors.

Procedure

The SART was the same as the semantic version used by McVay and Kane (2009), aside from the instructions. Subjects in the standard SART (replicating McVay & Kane, 2009) pressed the space bar as quickly as possible when frequent nontarget words appeared onscreen; they withheld responses to rare targets. Subjects in the vigilance SART, in contrast, did the reverse and withheld responses to frequent nontarget words and quickly pressed the space bar only when infrequent targets appeared.

Subjects completed 1,810 trials, with each stimulus centered for 250 ms, followed by a 950-ms mask of 12 capitalized Xs, the length of the longest stimulus. The first 10 (unanalyzed) buffer trials presented nontargets. The remaining trials comprised eight blocks, each presenting 225 trials consisting of 45 words repeating five times in a different random order. Within each set of 45, five targets appeared in random order among 40 nontargets (11% of trials). Thought probes followed 60% of targets within each block. After the first four blocks, subjects took a 30-s break.

Thought-probe screens presented the question “What were you just thinking about?” and seven response options. We instructed subjects to report what they were thinking in the moment before the probe appeared, and our instructions elaborated on these choices: (1) task: thinking about the stimulus words or appropriate response; (2) task performance: thoughts evaluating one's own performance; (3) everyday stuff: thinking about recent or impending life events or tasks; (4) current state of being: thinking about states such as hunger or sleepiness; (5) personal worries: thinking about life concerns, troubles, or fears; (6) daydreams: having fantasies disconnected from reality; (7) other: only for thoughts not fitting other categories. During the task, thought probes presented the response names (i.e., the italicized names just mentioned), and subjects pressed the corresponding number to indicate thought content.

Results

We report nondirectional null-hypothesis significance tests with alpha set to .05 and partial eta-squared (η^2) as an effect-size estimate.

SART performance

Accuracy

Mean accuracy rates for rare target trials were .42 in the standard SART (no-go trials) and .85 in the vigilance SART (go trials). For each subject, we calculated a signal-detection sensitivity score using the formula for logistic distributions (Snodgrass & Corwin, 1988), $d_L = \ln\{[H(1 - FA)]/[(1 - H)FA]\}$, and a C_L score, representing bias, using $C_L = 0.5[\ln\{[(1 - FA)(1 - H)]/[(H)(FA)]\}]$, where \ln = natural log, H = hit proportion, and FA = false-alarm proportion. We adjusted individual hit or false-alarm rates of 0 and 1 by .01. Negative C_L scores reflect a *go* bias. Table 1 presents d_L and C_L scores by SART and task block (quarters), where it appears that performance was better overall in the vigilance SART and that both d_L and C_L decreased over time only in the standard SART.

Table 1 is omitted from this formatted document.

A 2 (SART type) \times 4 (block) mixed-model analysis of variance (ANOVA) on d_L confirmed a main effect of block, $F(3, 828) = 13.02$, $MSE = 14.03$, $\eta_p^2 = .05$, and SART type, $F(1, 276) = 2,517.47$, $MSE = 3,494.82$, $\eta_p^2 = .90$, modified by an interaction, $F(3, 828) = 9.19$, $MSE = 9.89$, $\eta_p^2 = .03$. Sensitivity was higher in vigilance than in standard SART, and performance dropped significantly in the standard, $F(3, 423) = 28.52$, $MSE = 22.74$, $\eta_p^2 = .17$, but not the vigilance, $F(3, 423) < 1$, task. A 2 (SART type) \times 4 (block) mixed-model ANOVA on C_L indicated a main effect of block, $F(3, 828) = 4.08$, $MSE = 0.96$, $\eta_p^2 = .02$, and SART type, $F(1, 276) = 2,185.73$, $MSE = 2,854.80$, $\eta_p^2 = .89$, modified by an interaction, $F(3, 828) = 5.82$, $MSE = 1.36$, $\eta_p^2 = .02$, indicating a *go* bias in standard SART that decreased over blocks, $F(3, 405) = 11.43$, $MSE = 2.19$, $\eta_p^2 = .08$, and a *stableno-go* bias in vigilance SART, $F(3, 423) < 1$.

RT

The requirements of the two different SART versions produced different RT data for analysis: In standard SART, subjects responded to frequent nontarget trials, whereas in vigilance SART, subjects responded only to the infrequent target trials. Thus, RTs were based on a maximum of 1,600 trials per standard SART subject but only 200 trials per vigilance SART subject. RT data from the two SARTs are, therefore, not directly comparable, and we report them separately.

Standard SART

Table 1 presents the mean RT and mean intraindividual RT variability (standard deviation) on accurate nontarget (go) trials across blocks. A repeated-measures ANOVA on M RT indicated that subjects responded more quickly as the task progressed, $F(3, 423) = 13.03$, $MSE = 25,812.22$, $\eta_p^2 = .09$. Subjects' RTs to nontarget go trials also became more variable with time on task, $F(3, 423) = 45.27$, $MSE = 46,930.73$, $\eta_p^2 = .24$.

Standard SART RTs to nontargets were shorter preceding target errors than preceding accurate *no-go* responses (Cheyne et al., 2009; McVay & Kane, 2009; Robertson et al., 1997; Smallwood et al., 2004) and shorter preceding TUTs than on-task reports (McVay &

Kane, 2009). As in McVay and Kane (2009), we collapsed RTs over the four nontarget trials preceding a target, and as expected, they were shorter ($M = 418$) preceding errors than preceding correct *no-go* responses ($M = 502$), $t(141) = 23.81$. TUT reports were also preceded by faster RTs ($M = 442$) than were on-task thought reports ($M = 462$), $t(141) = 4.21$. Replicating McVay and Kane, the RT decrease preceding errors was considerably larger than that preceding TUT reports, suggesting that SART errors and TUT reports were not isomorphic.

Vigilance SART

Table 1 presents the means and intraindividual variability of RTs on accurate target (go) trials across quarters. A repeated-measures ANOVA on M RT indicated that, in contrast to subjects in the standard SART, subjects in the vigilance SART responded more slowly as the task progressed, $F(3, 423) = 16.64$, $MSE = 25,450.09$, $\eta_p^2 = .11$. Subjects' RTs to rare go target trials also became more variable with time on task, $F(3, 423) = 5.37$, $MSE = 3,775.90$, $\eta_p^2 = .04$.

Thought reports

In the standard SART, subjects reported task-related thoughts and task-unrelated thoughts (TUTs) on an average of 22.1% and 52.5% of the probes, respectively. We defined TUTs as thoughts about *current state* ($M = 22.4\%$), *daydreams* (10.4%), *everyday stuff* (8.6%), *worries* (4.9%), and *other* (6.4%). In the vigilance SART, subjects reported task-related and task-unrelated thoughts on 29.8% and 52.8% of the probes, respectively: *current state* ($M = 24.8\%$), *daydreams* (9.9%), *everyday stuff* (7.9%), *worries* (5.4%), and *other* (4.9%). The remaining probe responses represented self-evaluative thoughts, or “task-related interference” (TRI; Smallwood et al., 2006); because TRI represents a hybrid between task-related and task-unrelated thinking, that is, they are about one's task performance but not about the task itself, we later address it separately.

Figure 1 shows that TUT rates increased, and on-task thought decreased, over blocks in both SARTs. For TUTs, a 2 (SART type) \times 4 (block) mixed ANOVA indicated an increase over blocks, $F(3, 846) = 251.24$, $MSE = 5.09$, $\eta_p^2 = .47$, that didn't differ between SART types, $F(3, 846) = 1.83$, $\eta_p^2 = .01$, $p = .14$. In a parallel analysis, on-task thought rates decreased over blocks, $F(3, 846) = 90.73$, $MSE = 1.43$, $\eta_p^2 = .24$, and were lower for standard than for vigilance subjects, $F(1, 282) = 10.54$, $MSE = 1.67$, $\eta_p^2 = .04$; block and SART type did not interact, $F(3, 846) = 1.81$, $\eta_p^2 = .01$, $p = .14$.

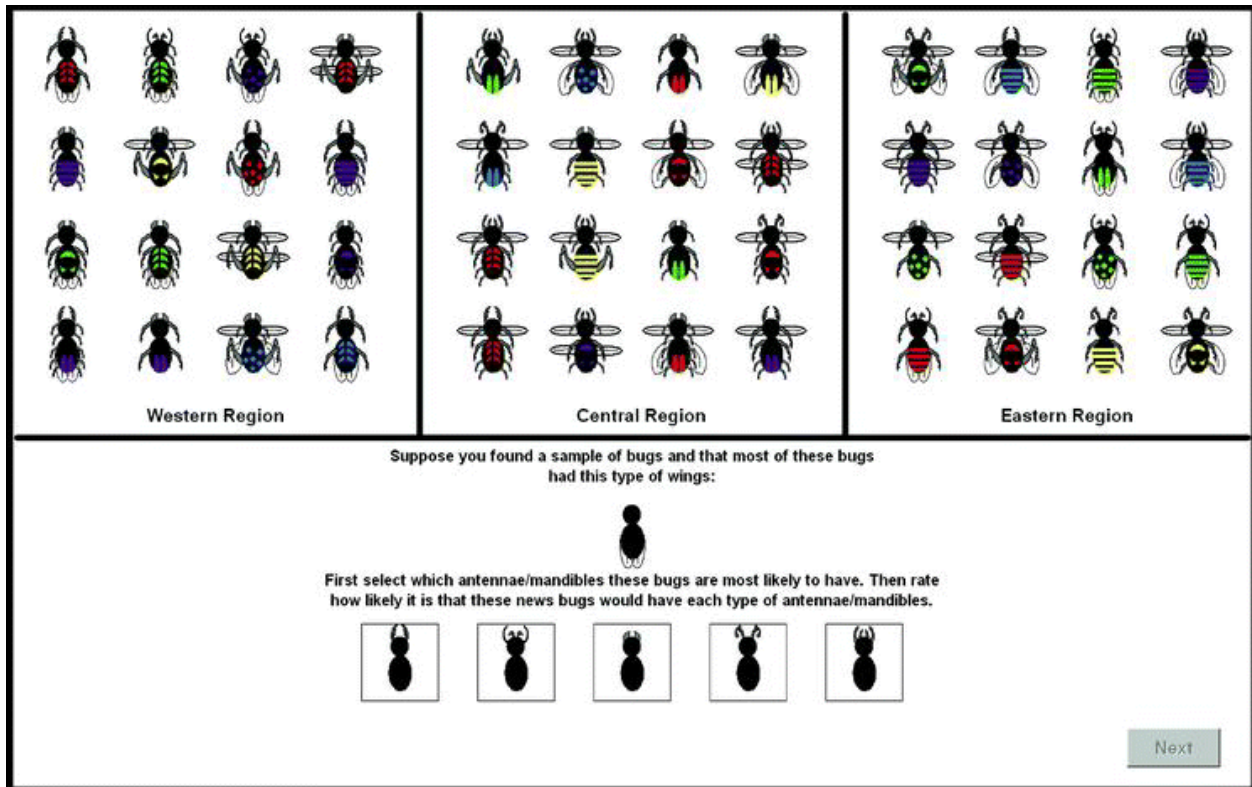


Figure 1. Mean proportion of thought reports, by SART type (standard, vigilance), by thought category (on-task, TUT), across task blocks ($N = 284$). Error bars represent standard errors. TUT = task-unrelated thought; OnTask = on-task thought; SART = sustained attention to response task.

Performance by thought report

In both SARTs, in-the-moment target accuracy was predicted by thought content (see Figure 2). A 2 (thought report) \times 2 (SART type) \times 4 (block) repeated-measures ANOVA indicated more accurate responding preceding on- versus off-task thought reports, $F(3, 161) = 217.65$, $MSE = 14.21$, $\eta_p^2 = .58$, and in the vigilance SART versus the standard SART, $F(1, 161) = 160.82$, $MSE = 35.94$, $\eta_p^2 = .50$; these main effects were modified by a Thought Report \times SART Type interaction, $F(1, 161) = 32.93$, $MSE = 2.15$, $\eta_p^2 = .17$ (no block effects or interactions were significant; $F_s < 2.00$ and $p_s > .10$). Although target accuracy was significantly higher for trials preceding on-task thought reports versus TUTs for both standard and vigilance SARTs ($t_s > 8.80$), this thought-content

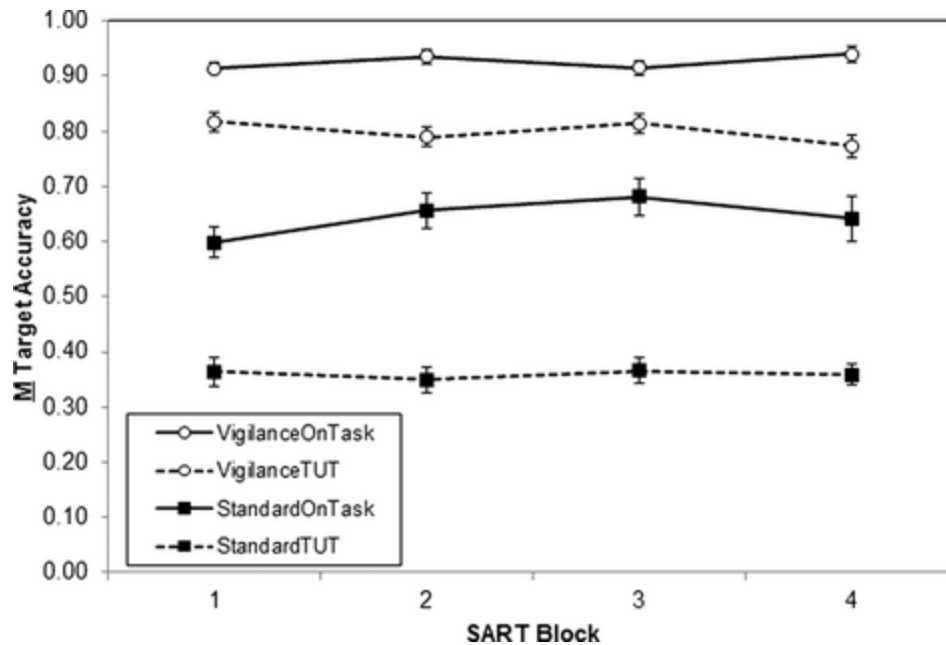


Figure 2. Mean accuracy on target trials, by SART type (standard, vigilance), by thought category (on-task, TUT), across task blocks ($N = 284$). Error bars represent standard errors. OnTask = on-task thought; TUT = task-unrelated thought; SART = sustained attention to response task.

At the level of intratask individual differences, subjects' TUT rate in standard SART significantly predicted d_L ($r = -.30$), mean nontarget RT ($r = -.17$), and nontarget RT variability ($r = .25$) but not C_L ($r = -.01$); in vigilance SART, TUT rate also predicted d_L ($r = -.31$) but not mean target RT ($r = -.01$), target RT variability ($r = .05$), or C_L ($r = .03$). TUT– d_L correlations increased significantly from Block 1 to Block 2 in both the standard SART ($r_s = -.08, -.34, -.32, \text{ and } -.35$ for Blocks 1–4, respectively) and vigilance SART ($r_s = -.06, -.36, -.31, \text{ and } -.25$), as indicated by Williams's t test (Steiger, 1980).

Intertask individual differences

WMC scores (z score composites) did not differ between SART groups (standard SART $M = 0.14$, $SD = 0.88$, and vigilance $M = 0.13$, $SD = 0.85$; $t < 1$). Table 2 presents the reliability and correlation coefficients for the critical variables for the standard and vigilance SARTs. As expected (and replicating McVay & Kane, 2009), WMC variation predicted performance and thought in the standard SART, correlating modestly but significantly with d_L , within-subject RT variability, and TUT rate. In contrast, in the vigilance SART, WMC variation did not predict performance or TUT rate, despite higher TUT rates predicting worse performance (d_L and RT variability), as noted previously.

Table 2 is omitted from this formatted document.

Table 3 presents the results from hierarchical regression analyses predicting standard SART accuracy (d_L) with WMC and TUT rate (we did not run parallel analyses on the vigilance SART because WMC did not predict performance). Replicating McVay and Kane (2009), WMC and TUT rate each accounted for shared and unique d_L variance. WMC accounted for 10.1% of the d_L variance, of which 2.9% (almost one third) was shared with TUT rate and 7.2% was unique; TUTs predicted 6.4% of the SART d_L variance independently of WMC (total $R^2 = .165$). Table 4 presents hierarchical regression analyses for standard SART RT variability, where WMC accounted for 13.4% of the variance, of which 2.7% (about one fifth) was shared with TUT rate and 10.7% was unique; TUT rate accounted for 3.7% of the variance beyond WMC (total $R^2 = .171$).

Table 4 is omitted from this formatted document.

Subjects' thoughts about their own performance (TRI) comprised 25.4% of responses in the standard SART (M proportions = .37, .25, .21, and .18 over blocks) and 17.4% of responses in the vigilance SART (M s = .24, .17, .15, and .14 over blocks). A 2 (SART type) \times 4 (block) mixed ANOVA indicated that TRI rates decreased across blocks, $F(3, 846) = 89.13$, $MSE = 1.14$, $\eta_p^2 = .24$, and that they were higher in the standard than vigilance SART, $F(1, 282) = 17.16$, $MSE = 1.82$, $\eta_p^2 = .06$. These main effects were also modified by an interaction, $F(3, 846) = 7.99$, $MSE = 0.10$, $\eta_p^2 = .03$. Although TRI decreased over blocks in both conditions, the Block \times SART Type interaction was significant from Block 1 to Block 2, indicating a steeper decrease in TRI reports in the standard than in the vigilance SART, $F(1, 282) = 7.02$, $MSE = 0.18$, $\eta_p^2 = .02$.

Like TUTs, instances of TRI tended to predict in-the-moment errors in both tasks. Target accuracy rates were lower following TRI than following on-task thoughts in the standard SART (M s = .38 vs. .62 for TRI vs. on-task), $t(134) = 10.54$, and the vigilance SART (M s = .86 and .93 for TRI vs. on-task), $t(136) = 5.10$. Regarding individual differences, however, TRI behaved differently from TUTs: TRI rate did not significantly predict d_L , mean RT, or RT variability in either SART (r s = .02, -.06, and -.04 for standard; .15, .08, and -.11 for vigilance). Moreover, WMC did not predict TRI rate in either the standard or vigilance SART (r s = -.04 and -.05, respectively). 5

Discussion

The vigilance version of the SART yielded better performance than did the standard SART but the same rate of mind wandering. In both versions, TUTs increased with time on task and predicted (if not affected) performance on target trials. However, in the moment, TUTs seemed more detrimental to performance in the standard than the vigilance SART. In support of executive attention theory (Engle & Kane, 2004), WMC predicted TUTs in the standard SART. Moreover, TUTs and WMC each accounted for shared and unique variance in SART performance, highlighting the role of goal maintenance as one component of attention control

that varies with WMC. In the vigilance SART, in contrast, WMC did not predict TUTs or performance. Although we did not predict this null effect, it is consistent with previous findings where WMC only selectively predicted mind wandering or performance, during demanding tasks that require a particular form of executive control (Kane, Brown, et al., 2007; Poole & Kane, 2009; see General Discussion for a more thorough treatment). We designed the vigilance SART to reduce the need for competition resolution and, thereby, the variance predicted uniquely by WMC, beyond that shared with TUT rate. Instead, we have identified a task where TUTs significantly affect performance but WMC does not predict either thought or performance.

Lapses of Attention and RT Distributions

Here we evaluated the relation between WMC and subjects' longest RTs using both the ranking and binning method (for a review, see Coyle, 2003) and individualized ex-Gaussian distributions (Schmiedek et al., 2007; Unsworth et al., 2010). Furthermore, we used TUTs, the subjective experience accompanying attentional lapses, as an initial means to test whether the worst performance rule is best explained without referring to attentional lapses, the conclusion drawn by Schmiedek et al. (2007) on the basis of RT modeling and parsimony. That is, they were able to account for WMC-related individual differences in long RTs using a quantitative model that had no parameter reflecting attentional lapses.

The SART is unique among tasks used to investigate the worst performance rule, however, in that attentional lapses are hypothesized to produce occasional very fast responses in addition to occasional very slow responses. That is, unlike in the choice-RT tasks that researchers typically examine (including Schmiedek et al., 2007), the frequency of nontarget trials in the SART builds a habitual go response. This prepotency of the go response makes the shortest RTs interesting as well, because excessively short RTs may reflect responses emitted before stimulus analysis is complete, as a result of habitual, mindless responding (Cheyne et al., 2009). We therefore expected subjects' shortest RTs, as well as their longest RTs, to be related to their mind-wandering rates.

Method

Subjects

We combined data from McVay and Kane (2009; $N = 244$) and from the standard SART in the present experiment ($N = 142$) for a total of 386 subjects.

SART versions

The SART from McVay and Kane (2009) was the same as the present standard SART, with one exception. In McVay and Kane, subjects completed a perceptual, perceptual–semantic, or semantic version (the latter was identical to the present standard task). The perceptual and perceptual–semantic SARTs instructed subjects to respond to lowercase words and withhold

responses to infrequent uppercase words; in the perceptual–semantic SART, letter case also predicted perfectly the semantic category (e.g., all uppercase words were animal names and lowercase words were food names). McVay and Kane reported no differences in TUT rate or performance (dL) between the three SART types, and so we here combined data from all SART versions.

Analyses

For each subject, we fitted an ex-Gaussian function using quantile maximum-likelihood estimation (QMLE; Brown & Heathcote, 2003) to nontarget RTs. Due to program limitations, RTs from only 1,199 trials from each subject can be entered. The SART has the potential for 1,600 nontarget RTs, and so we trimmed the data set for analysis. We first trimmed ambiguous RTs, those that may have been late responses from a previous trial or anticipations to the current trial ($RT < 150$ ms). We then fitted the distributions twice, once using 1,199 randomly selected trials for each subject and once using all 1,200 trials from the second, third, and fourth SART blocks (recall that these blocks yielded much higher TUT rates than did Block 1, as well as significantly higher correlations between TUT rate and performance).

Results and Discussion

For a visual representation of WMC differences in RT variability (see also Unsworth et al., 2010), Figures 3A and 3B present 100 randomly selected RTs for two randomly selected higher WMC subjects (composite z scores = 1.53 and 1.50) and for two randomly selected lower WMC subjects (z scores = -2.15 and -2.24). Figure 4 presents ranked RTs from the same subjects. The low WMC subjects show much greater RT variability from trial to trial and a greater range of RTs within the task.

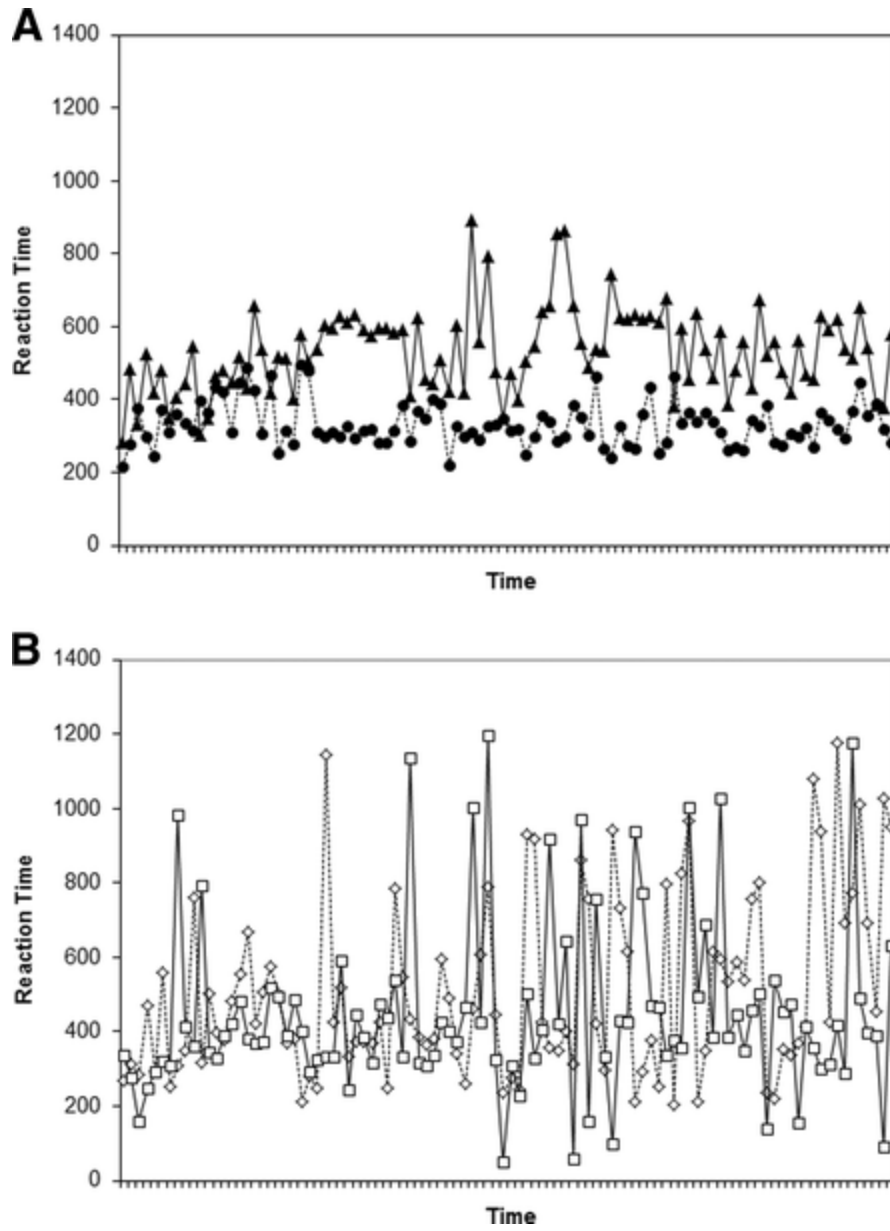


Figure 3. Reaction times (RTs) in milliseconds for 100 randomly selected trials for two randomly selected higher working memory capacity (WMC) subjects (Panel A) and two randomly selected lower WMC subjects (Panel B) from the standard sustained attention to response task.

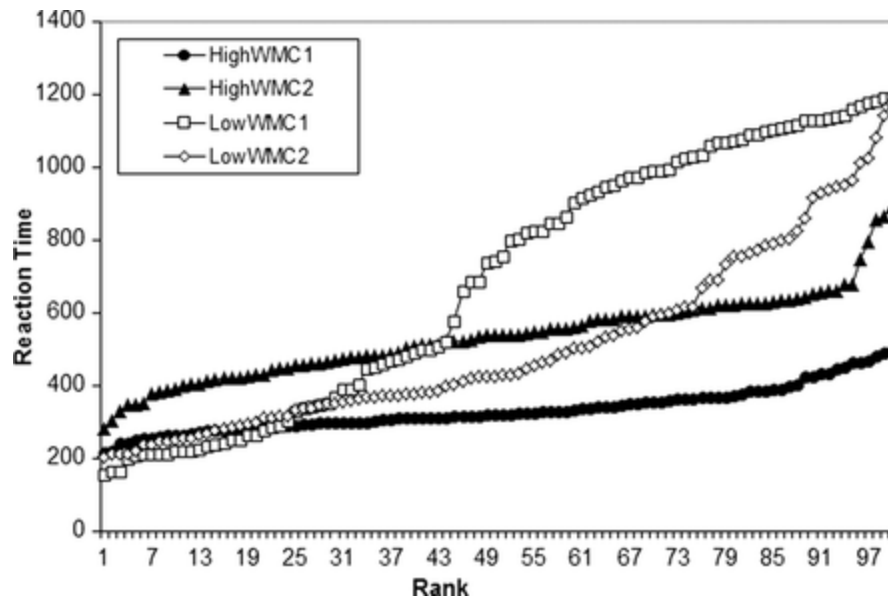


Figure 4. Ranked reaction times (RTs) in milliseconds for 100 randomly selected trials for two randomly selected higher working memory capacity (WMC) subjects and two randomly selected lower WMC subjects from the standard sustained attention to response task. HighWMC = higher WMC subject; LowWMC = lower WMC subject.

Descriptive statistics, for the ranked RTs by quintile, and the ex-Gaussian parameter estimates, are shown in Table 5. The theoretically critical τ parameter did not differ when estimated from 1,199 random cases ($M = 117$) or from trial Blocks 2–4 of the standard SART ($M = 117$), $t(384) < 1$, so we simply report the analyses on RT data from Blocks 2–4, which is a preferable strategy because TUTs and performance measures were all drawn from the same set of trials.

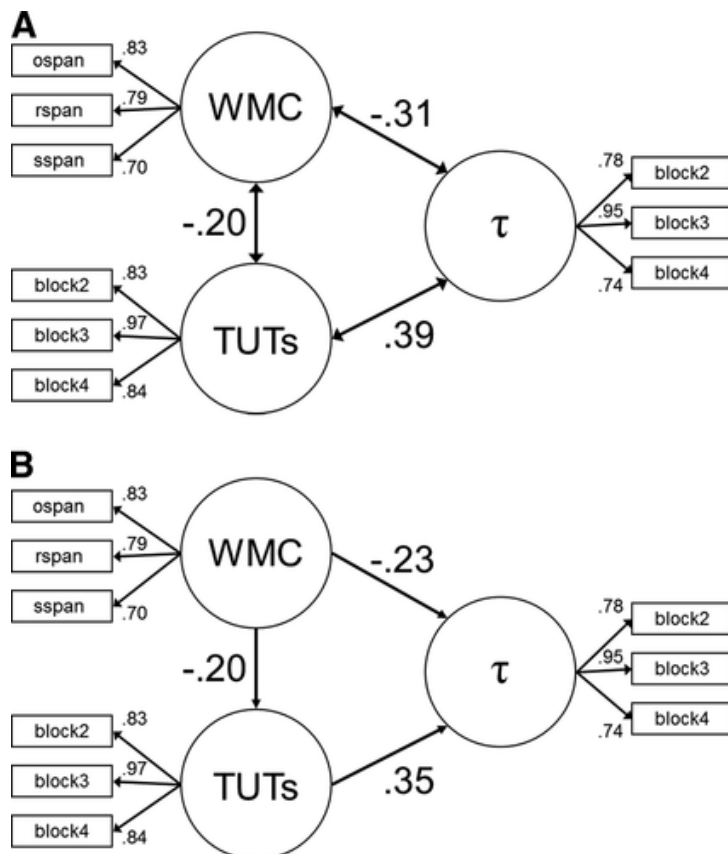
Table 5 is omitted from this formatted document.

Table 6 shows the correlations between the RT quintiles and WMC, TUT rates, and SART accuracy (dL). WMC correlated negatively with the longest RTs and positively with the shortest, indicating that higher WMC subjects had more consistent and moderate RTs (i.e., their fastest RTs, in Quintile 1, were relatively long, and their slowest RTs, in Quintile 5, were relatively short) than did lower WMC subjects. TUT rate showed the inverse (but logically consistent) pattern to WMC, with positive correlations with the slowest quintile and negative correlations with the fastest. Like people with lower WMC, people with higher TUT rates responded more variably, with their shortest times being shorter and longest times being longer than those of people with lower TUT rates. Regarding SART accuracy, the gradual change from positive to negative correlations across RT quintiles indicates there was an optimal, intermediate response speed for accurate performance on the SART: Subjects with the shortest RTs performed poorly, but so did those with the longest.

Table 6 is omitted from this formatted document.

The relations among WMC, TUT, standard SART performance, and the ex-Gaussian RT parameter estimates (see Table 6) tell a story similar to that of the longest RTs from the quintile analyses. As predicted, WMC and TUT rate predicted the τ parameter, indicating that subjects with lower WMC or higher TUT rates had more positively skewed RT distributions, reflecting their more frequent slowed responses. In slight contrast, dL correlated positively with μ and negatively with τ , indicating that slow-but-steady RTs yielded the most accurate performance in the SART. Of note, the correlations among our individual measures of primary interest—WMC, TUT rate, and τ —were not very large, and they were considerably weaker than those among the latent variables reported by Schmiedek et al. (2007).

For a better parallel to their analysis, then, we used our multiple WMC indices and multiple blocks of the SART task to derive latent variables for WMC, TUT, and τ . Figure 5A depicts a confirmatory factor analysis on these three constructs (using Blocks 2–4 of the SART to derive TUT rate and τ latent variables), and the model fit the data well, $\chi^2(24) = 41.91$, $\chi^2/df = 1.75$, comparative fit index (CFI) = .99, root-mean-square error of approximation (RMSEA) = .044, standardized root-mean-square residual (SRMR) = .028. Here, WMC and TUT rate each correlated substantially with τ , and the magnitude of the WMC and TUT correlations with τ were more similar here than in the univariate analyses reported in Table 6.



*Figure 5. Latent variable analyses testing the relations among working memory capacity (WMC), task-unrelated thought (TUT) rate (measured across sustained attention to response task [SART] Blocks 2–4), and the reaction time parameter τ (measured across SART Blocks 2–4). Circles represent latent variables, and square boxes represent observed variables. Panel A: Confirmatory factor analysis; double-headed arrows connecting latent variables to each other represent the correlations between the constructs, and numbers appearing next to each single-headed arrow represent the loadings for each manifest variable onto the latent variable. Panel B: Structural equation model with TUT rate as a partial mediator of the WMC– τ association; single-headed arrows connecting latent variables with each other are analogous to semipartial correlations between these constructs. All depicted path coefficients are statistically significant. *ospan* = operation span; *rspan* = reading span; *sspan* = symmetry span.*

Regression analyses and a structural equation model indicated that WMC and TUT rate accounted for both shared and unique variance in the RT indices. Table 7 presents the coefficients for a hierarchical regression analysis predicting the slowest RT quintile, the fastest quintile, and the τ parameter with WMC and TUT rate. WMC explained significant variance in the slowest RT quintile ($R^2 = .06$), here with one third of that variance shared with TUT rate. WMC accounted for significant variance in the fastest RT quintile as well ($R^2 = .04$), here with almost half of that variance shared with TUT rate. Finally, the significant but modest relation between WMC and τ ($R^2 = .03$) was substantially mediated by TUT rate, with over half of the WMC-related variance explained by mind wandering; TUT rate additionally predicted considerable variance independently of WMC (total $R^2 = .11$).

Table 7 is omitted from this formatted document.

As a conceptual replication of our regression analyses, here using latent-variable methods, Figure 5B presents a structural equation model that tested not only for the independent, direct effects of WMC and TUT rate on τ but also for the indirect, mediated effect of WMC on τ , through TUT rate. This partial-mediation model had identical fit statistics to our previous confirmatory factor analysis (from Figure 5A), here with WMC and TUT rate together accounting for 21% of the variance in τ . Of central importance, the WMC \times TUT indirect effect was $-.07$ ($p < .01$), indicating that some of WMC's prediction of τ was mediated by TUT rate (note that the full mediation model, in which WMC had no direct effect on τ , did not fit the data as well as did the partial-mediation model, as indicated by a significant χ^2 test, χ^2 difference $2(1) = 16.35$).

In summary, and consistent with the worst performance rule, subjects' WMC scores were negatively related to their longest RTs, calculated using both RT quintiles and individualized ex-Gaussian RT distributions. Mind wandering was positively related to long RTs, suggesting that these RTs might reflect, in part, lapses of attention experienced as off-task thoughts. Furthermore, TUT rate partially mediated the relationship between WMC and subjects' longest RTs. This analysis counters the claim of Schmiedek et al. (2007) that lapses of attention to the ongoing task do not contribute to the τ parameter estimates from individuals' ex-Gaussian

distributions. By actually measuring TUT experiences, we found a contribution of attentional lapses to the association between WMC and τ .

Lapses of Attention and Evidence-Accumulation Modeling of RT

Of course, the Schmiedek et al. (2007) approach to the question of WMC's association with long RTs had many methodological strengths, not the least of which was using a rigorous, quantitative model of RTs, the EZdiffusion model (Wagenmakers et al., 2007), which allowed them to test an attentional lapse theory of the worst performance rule without relying on subjective reports. As we argued earlier, however, their use of a reduced diffusion model—which had no parameter to reflect attentional lapses—weakened their argument from parsimony (i.e., from a null effect). Here, then, we used a more complete evidence-accumulation model, which includes a parameter for within-subject drift rate variability, to more directly test whether WMC predicts RT variability in part because it predicts lapses and TUTs.

Recent advances have simplified the execution of quantitative RT modeling (e.g., Vandekerckhove & Tuerlinckx, 2007, 2008; Voss & Voss, 2007; Wagenmakers et al., 2007). Here, we took advantage of a Microsoft Excel application (Donkin, Averell, Brown, & Heathcote, 2009) to implement the linear-ballistic accumulator (LBA) model (Brown & Heathcote, 2008). The LBA mathematically simplifies the drift process by eliminating within-trial stochasticity from the Ratcliff model (evidence accumulation is thus “ballistic”; e.g., Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004). Moreover, in contrast to the EZdiffusion model used by Schmiedek et al. (2007), the LBA takes as input the full distribution of subjects' correct and incorrect RTs, it yields a similar complement of parameters from choice-RT data to the Ratcliff model (including drift rate and drift-rate variability), it successfully accounts for a breadth of choice-RT-task phenomena similar to that of the Ratcliff model (Brown & Heathcote, 2008), and it yields parameter values similar to those from Ratcliff's model when they are both applied to the same data (Donkin, Brown, Heathcote, & Wagenmakers, 2011).

Quantitative evidence-accumulation models are typically applied to tasks requiring choice between two overt responses, such as word versus nonword in lexical decision, rather than to go/no-go tasks (such as the SART) with only one overt response. Recent work, however, has suggested that two-choice and go/no-go versions of the same task (whether lexical decision, numerosity discrimination, or recognition memory) can best be modeled by assuming an implicit decision boundary for no-go “responses” and by fixing drift rate and drift-rate variability to be equal across both two-choice and go/no-go task types, while allowing response bias, response criterion, and nondecision parameters to vary between task types (Gomez, Ratcliff, & Perea, 2007). Because two-choice and go/no-go tasks appear to engage identical evidence-accumulation processes, we felt justified in applying a quantitative evidence-accumulation model to our SART data, especially given that our central predictions hinged on the drift and drift-variability parameters.

Method

Subjects and SART versions

We used the same data set here as we did for the previous RT quantile and ex-Gaussian analyses.

Analyses

For each subject, we fitted all correct and error RTs for SART Blocks 2–4 (screened as they had been for the RT quantile and ex-Gaussian analyses) using the Donkin et al. (2009) Excel program. Our final data set included 374 subjects, each of whom had at least 750 RTs to model and whose RTs could be reliably fitted by Excel's Solver function (for some subjects, the function would settle on a local maximum rather than a global, optimal maximum; here, seeding the model with different starting parameter values yielded vastly different resulting parameter estimates, indicating invalid solutions). The LBA analyses produced estimates of the following parameters for each subject: drift rate (v_1 , in the Donkin et al., 2009, LBA Excel program; corresponding to the diffusion model's μ), drift-rate variability (s ; corresponding to the diffusion model's η), and nondecision processes (t_0 ; corresponding to T_{er}); we estimated response criterion, or cautiousness (in the diffusion model, a), by combining the a and b parameters from the LBA Excel program ($b - [a/2]$; see Donkin et al., 2011).

Results and Discussion

Mean values for the LBA parameters in this sample were .847 (SD = .210) for drift rate, .204 (SD = .118) for drift-rate variability, 76.33 (SD = 77.47) for nondecision time, and 302.03 (SD = 154.82) for response criterion. As is sometimes the case (e.g., Ratcliff & Tuerlinckx, 2002; but see Ratcliff, Thapar, & McKoon, 2010), correlations among these parameters were statistically significant, ranging from $r = .114$ (Drift Rate \times Nondecision Time) to $r = .635$ (Drift Rate \times Response Criterion); drift rate correlated with drift variability at $r = .411$, indicating substantial shared variance (including some likely statistical dependency from the data-fitting process) but not redundancy. Table 8 presents the correlations between these LBA parameters and our other variables of primary interest: WMC, TUT rate, and SART τ . Of importance, and consistent with the notion that no behavioral measure is a process-pure reflection of any one underlying mechanism, both performance and thought measures from the SART (TUTs and τ , respectively) correlated with multiple LBA parameters: τ with drift variability, nondecision time, and response criterion, and TUT rate with drift variability, nondecision time, and drift rate. However, as predicted, TUTs were most strongly correlated with drift-rate variability (higher TUT rate = greater drift-rate variability). WMC also correlated more strongly (in a negative direction) with drift-rate variability than with any other LBA parameter, and only its correlation with drift variability was statistically significant.

Table 8 is omitted from this formatted document.

Hierarchical regression analyses, presented in Table 9, tested whether drift rate or drift-rate variability would partially mediate the associations between WMC and TUT rate and between WMC and τ . We had predicted that drift-rate variability would be at least as strong a mediator as would drift rate, and indeed, given the near-zero correlation between WMC and drift rate ($r = .016$) it was unlikely that drift rate would be a stronger mediator than would drift-rate variability. As shown in Table 9, regarding TUT rate, the 3.6% of TUT variance predicted by WMC was essentially independent of drift rate, but nearly 40% of it was shared with drift-rate variability. Likewise, for τ , the 2.5% of τ variance predicted by WMC was independent of drift rate, but approximately 60% of it was shared with drift-rate variability. In short, WMC predicts TUTs and τ , in large part via its shared variance with drift-rate variability.

Table 9 is omitted from this formatted document.

As a further test of whether attentional lapses might contribute to the correlation between WMC and τ , we again supplemented our regression analyses with a latent variable analysis. Figure 6 presents a structural equation model that tested for mediation of the WMC– τ association by two variables that reflect, to some degree, attentional lapses: subjects' overall TUT rate (based on subjective self-report) and their drift-rate variability parameter estimate (based on the quantitative LBA model). The model provided a just-adequate fit to the data, $\chi^2(16) = 51.78$, $\chi^2/df = 3.24$, CFI = .97, RMSEA = .076, SRMR = .028. The WMC factor was modeled as the variance common to the three complex span tasks, and τ factor was modeled as the variance common to τ estimates from SART Blocks 2–4. We did not model an “attentional lapse” factor based on the shared variance between TUT rate and drift-rate variability because, even though they correlated significantly, it is inadvisable to model latent factors with fewer than three observed measures (Kline, 2005; moreover, drift-rate variability, but not TUT rate, was derived from the same RT data as the τ dependent variable). Our model did allow TUT rate and drift variability to correlate, however. As indicated in Figure 6, TUT rate and drift-rate variability both acted as partial mediators of the WMC– τ association, with the full complement of predictor variables accounting for 58% of the variance in τ . Specifically, the WMC \times TUT Rate indirect path was $-.04$ ($p < .01$), and the WMC \times Drift-Rate Variability indirect path was $-.10$ ($p < .01$). Some of WMC's prediction of τ , then, resulted from its influence on the experience of attentional lapses. Note, though, that this mediation was only partial, as the model in Figure 6 fitted the data significantly better than did one that eliminated the direct path from WMC to τ , $\chi^2_{\text{difference}}(1) = 18.71$.

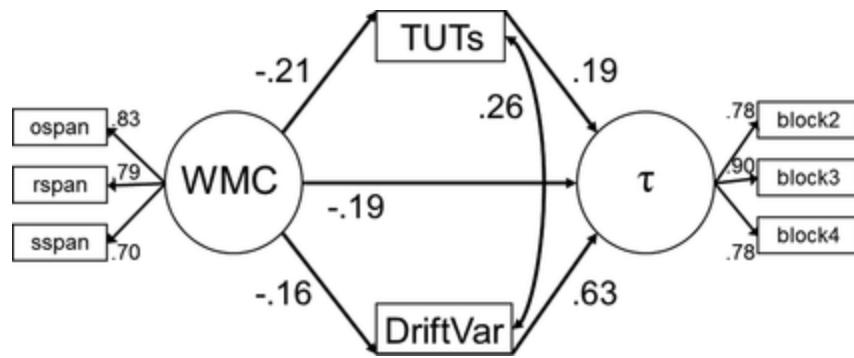


Figure 6. Structural equation model testing for mediation of the WMC– τ association by two indices of attentional lapses: TUT rate and DriftVar. Circles represent latent variables, and square boxes represent observed variables. All depicted path coefficients are statistically significant. TUTs = task-unrelated thoughts; WMC = working memory capacity; τ = mean and standard deviation of the exponential component of the ex-Gaussian distribution from the sustained attention to response task; DriftVar = drift-rate variability parameter from the linear-ballistic accumulator model; Block 2–Block 4 = sustained attention to response task Blocks 2–4.

In summary, our primary conclusions from quantitative RT modeling (using the Brown-Heathcote, 2008, LBA model) of a long-duration go/no-go task is that individual differences in τ are correlated more strongly with drift-rate variability than with drift rate (or with other LBA parameters). That is, as predicted from an attentional lapse framework, the extent to which subjects showed more skewing of their RT distribution (indicative of occasional, especially slow responses) seems to be driven to some significant degree by the extent to which subjects also showed increased trial-to-trial variability in the rate at which they accumulated response-relevant information from imperative stimuli. We suggest, along with Schmiedek et al. (2007), that one potential source of drift-rate variability is the experience of occasional attentional lapses. Of course, such lapses may not be the only source of intrasubject variation in drift rate, but their contribution is supported by our TUT findings, as well. That is, TUT rate correlated more strongly with drift-rate variability than with any other LBA parameter (including drift rate), indicating that people who had the most variable drift rates also experienced the most mind-wandering episodes during the SART. Moreover, the modest association between WMC and τ was substantially—and similarly—mediated by both drift-rate variability and TUT rate. In contrast to Schmiedek et al. (2007), then, we found evidence that attentional lapses make some contribution to τ , to the worst performance rule, and to WMC-related variation in worst performance.

Principal Components Analyses of RT Series

Using the combined standard SART data from the current experiment and McVay and Kane (2009), we conducted a principal components analysis (PCA) to identify intraindividual patterns of RT change. Our goal, like that of Smallwood, McSpadden, et al. (2008), was to seek RT patterns that signaled, in advance, the imminent commission of no-go errors, the experience of

TUTs, or both. Recall that Smallwood, McSpadden, et al. found that RT patterns identified through this PCA method predicted both SART errors and TUTs: A change from slow to quick responding in the trials leading up to a target or probe predicted errors, a change from quick to slow responding predicted on-task thoughts, and generally fast responding across the entire series predicted TUTs occurring without awareness. Here we attempted to replicate their findings while improving upon their methods. Specifically, we standardized RTs within-subject (with a z score transformation), used 20 trials prior to targets and probes, and compared errors to correct target trials and TUTs to on-task thoughts, rather than to potentially contaminated baselines (as in Smallwood, McSpadden, et al., 2008). We also used hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) to evaluate differences in the RT patterns leading up to error or TUT trials. The data have a hierarchical structure in which runs of nontarget trials (Level 1 data) are nested within-subject (Level 2 data) and are therefore best evaluated with a multilevel approach such as HLM.

Method

Smallwood, McSpadden, et al. (2008) conducted PCA on their SART RT data from runs of 12 trials that preceded each target stimulus or thought probe. Our PCA analyses assessed the 20 nontarget trials preceding every target trial, using the combined data set from McVay and Kane (2009; $N = 244$) and the present standard SART sample ($N = 142$). The PCA treated each set of 20 trials leading up to a target trial as one data series, so the N s we report later represent the number of series, not subjects. Prior to analysis, we excluded nontarget error trials, and we standardized RTs for each subject (against that subject's mean RT for the experiment, thus expressing each RT as a z score) to eliminate between-subjects RT differences that might mask or distort within-subject RT changes, as they may have in the Smallwood, McSpadden, et al. data (where RTs were not standardized). Targets were randomly presented in the SART, and so many targets had fewer than 20 trials between their appearances; these were not included. PCA analyzes only series without missing values, resulting here in 3,733 series of 20 trials before the appearance of a target.

Results and Discussion

Our first PCA (without rotation) yielded four components with eigenvalues greater than 1, accounting for 65.1% of the total variance; inspection of the scree plot, however, suggested that three components might be more appropriate and also consistent with the PCA solution from Smallwood, McSpadden, et al. (2008). We therefore reconducted the analysis to yield three unrotated components, which together accounted for 59.8% of the variance and which are presented in Table 10 (where “n back” refers to the distance from the target trial, from 20 back to 1 back, with 1 back representing the nontarget trial immediately preceding the target).

Table 10 is omitted from this formatted document.

Component 1 represents a general RT component, characterized by uniformly positive loadings across all n-back trials and thus indicating that different runs of 20 trials deviated, as a whole, from a subject's average RT (with some runs generally faster than average and others generally slower than average). Component 2 is characterized by inverse loadings for the beginning and end of the series (i.e., strong positive loadings among trials further back in the series and strong negative loadings among trials closer to the target), reflecting trial runs that were characterized by linear change (a speedup or slowdown) preceding the target event. Component 3 reflects series with a quadratic pattern of RTs across n-back trials (i.e., RTs getting relatively long and then short as the target approached, or vice versa).

The PCA yielded a score on each of the three components for each of the 3,733 RT series. This component score represented the extent to which each RT series matched the pattern expressed by the loadings presented in Table 10. For Component 1, RT series with positive scores were slower than the subject's average, and series with negative scores were faster than average. For Component 2, positive scores reflected a speedup as the target approached, and negative scores reflected a slowdown; scores of larger absolute magnitude indicated a steeper slope. For Component 3, RT series with positive scores started off slower, got faster, and then got slower as the target approached, and series with negative scores started off faster, got slower, and then got faster; again, larger absolute-magnitude scores represented steeper changes over pretarget trials.

We then used HLM to examine the PCA component scores as predictors of TUTs and target accuracy. Both TUTs and target accuracy were dichotomous variables (on-task vs. TUT; error vs. correct), so we used a Bernoulli distribution to evaluate the effects. We first tested whether any of these RT patterns predicted no-go errors. Figure 7 shows that the mean scores for all three components differed as a function of whether the RT series resulted in a commission error or an accurate no-go response to the target. The mean score of Component 1 was significantly higher for accurate trials than for error trials ($b = 0.786$, $SE = 0.045$), $t(3729) = 17.453$, $p < .001$, suggesting that subjects' relatively slower series were followed by better performance and that their faster series were followed by poorer performance. The mean score for Component 2 was significantly higher for error trials than for accurate target trials ($b = -0.393$, $SE = 0.040$), $t(3729) = -9.780$, $p < .001$: When subjects sped up more across trials before a target, they were more likely to make a commission error, whereas if they slowed down, they were more likely to correctly withhold their response. Error trials also had significantly lower scores on Component 3 than did accurate trials ($b = 0.254$, $SE = 0.040$), $t(3729) = 6.326$, $p < .001$, indicating that when subjects reacted faster at the beginning and end of the series, as opposed to in the middle, they were more likely to commit an error on the subsequent target (again indicative that a speedup just before a target predicted a commission error).

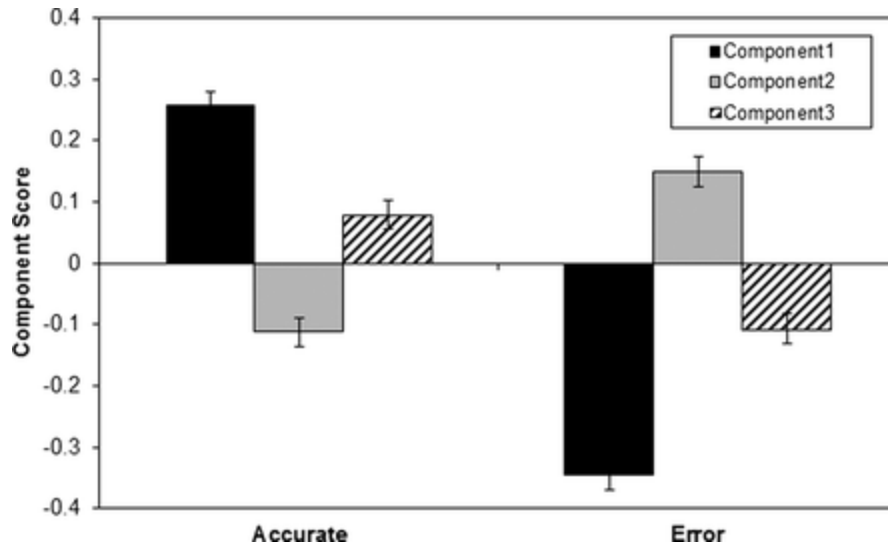


Figure 7. Component scores, by target accuracy, calculated using three-component principal components analysis on accurate nontarget reaction time sequences in the standard sustained attention to response task ($N_{series} = 3,427$). Error bars represent standard errors.

We next tested whether any of the RT patterns predicted thought content. Analyses of thought type were based on fewer series ($N = 2,927$) because not all target trials were followed by thought probes. Here, only the general RT component predicted TUTs (see Figure 8). The mean score for Component 1 was significantly higher for reports of on-task thinking than for TUTs ($b = 0.161$, $SE = 0.060$), $t(2723) = -4.26$, $p < .001$, indicating that series with generally longer RTs than a subject's average predicted on-task thoughts. Neither Component 2 scores ($b = -0.187$, $SE = 0.040$), $t(2723) = 1.303$, $p = .193$, nor Component 3 scores ($b = 0.007$, $SE = 0.038$), $t(2723) < 1$, $p = .853$, predicted TUTs.

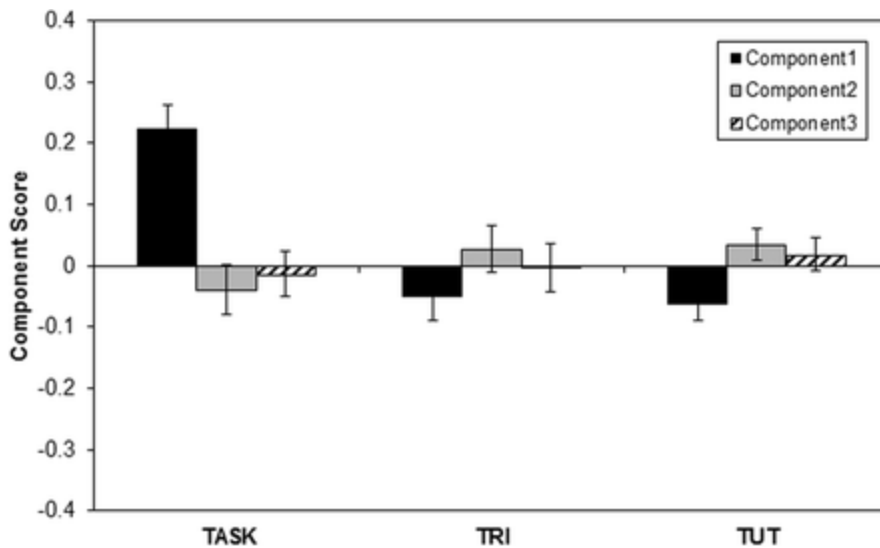


Figure 8. Component scores, by thought report, calculated using three-component principal components analysis on accurate nontarget reaction time sequences in the standard sustained attention to response task (Nseries = 2,693). Error bars represent standard errors. TASK = on-task thought; TRI = task-related interference; TUT = task-unrelated thought.

In summary, and consistent with Smallwood, McSpadden, et al. (2008), the three RT patterns revealed by our PCA (general relative speed, linear RT change, and quadratic RT change) predicted SART errors: Subjects were more likely to make an error on a target trial following generally faster-than-average nontarget RT series, as well as following nontarget RT series that increased in speed just prior to the target event (either as a linear change or as the end of a quadratic change, with slower RTs during the middle of the series). Only one of the three time-series patterns, however, predicted TUTs. Subjects were more likely to report an on-task thought, as opposed to a TUT, following a series of consistently slower-than-average RTs. In contrast to the findings from Smallwood, McSpadden, et al., changes in nontarget RTs preceding a target trial (whether linear or quadratic) did not reliably predict off-task thinking. We cannot be certain why we did not replicate their finding that decreasing pretarget RTs precede TUTs, but it may be because half of their trials came from a SART with a much slower rate of stimulus presentation than that used here (2,500 ms vs. 1,250 ms). Perhaps with very slow stimulus trains, RT variation is more reflective of thought content, whereas with faster trains, RT variation is more indicative of speed-accuracy tradeoff or yet some other mechanisms.

General Discussion

The three analytic sections of this article assessed (1) WMC-related differences in the performance of a standard (go/no-go) versus a vigilance (no-go/go) version of the SART, as well as the contributions of TUT experiences to each; (2) the role, if any, of attentional lapses in producing the worst performance rule and the association between WMC and subjects' longest RTs; and (3) the potential for within-subject RT patterns to objectively mark or predict TUT experiences online. In the original experiment presented here, we replicated the McVay and Kane (2009) finding that individual differences in WMC predict both TUT rates and performance (accuracy and RT variation) in the standard SART. However, WMC differences did not emerge in either performance or thought content from the vigilance SART. We then demonstrated (with the present data combined with those from McVay & Kane, 2009) that both TUT rate and the drift-rate-variability parameter (from quantitative evidence-accumulation choice-RT modeling) accounted for significant variance in the τ parameter of individual RT distributions from the standard SART. Moreover, these subjective (TUT rate) and objective (drift-rate variability) indices of attentional lapses partially mediated the association between WMC and τ . Finally, with our combined data set, we partially replicated the Smallwood, McSpadden, et al. (2008) findings that some within-subject RT patterns predict some measures of interest from the standard SART: Slower-than-average RT series predicted accurate responses and on-task thought reports, but dynamic RT changes preceding targets predicted only accuracy and not thought content.

Mind Wandering and Theoretical Accounts of WMC

A primary goal of this study was to leverage the probed thought report as a tool to assess attentional lapses as TUT experiences and, thereby, to evaluate theories of WMC variation, in particular the executive attention theory (e.g., Engle & Kane, 2004; Kane, Conway, et al., 2007; Unsworth & Spillers, 2010) and coordinative binding theory (e.g., Oberauer et al., 2007; Wilhelm & Oberauer, 2006). Whereas executive attention theory argues for the importance of attention-driven goal maintenance (and, as its failure, goal neglect) to WMC variation and its predictive power, coordinative binding theory suggests that such attention-control processes are only weakly or spuriously associated with WMC and that WMC's covariation with other constructs reflects a (nonexecutive) capacity to simultaneously bind multiple independent mental representations, such as stimulus–response (S–R) associations. Our finding that variation in TUT rate contributed to the WMC \times Standard SART correlation replicated McVay and Kane (2009) and supported (along with our drift-variability findings) the executive attention theory by demonstrating a contribution of attentional lapses to WMC's prediction of task performance. At the same time, this finding seems inconsistent with the variety of theories denying a role for attention control in WMC variation or covariation with other constructs (e.g., Colom et al., 2008; Mogle et al., 2008; Oberauer et al., 2007).

Standard versus vigilance SART

We designed the present study to go further, however, in explaining WMC's association to the standard SART—and, by extension, WMC's relation to other tasks requiring restraint of habitual responses in favor of novel ones (e.g., Long & Prat, 2002; Unsworth et al., 2004). Both here and in McVay and Kane (2009), failures of goal maintenance could not completely explain the WMC–performance correlations, because WMC was only modestly associated with TUT rate and, furthermore, TUT rate only partially mediated WMC's effects. That is, WMC consistently predicted significant variance in SART accuracy and RT variation independently of mind-wandering rate. McVay and Kane explained these findings by appealing to dual-process conceptions of executive control (e.g., Braver et al., 2007; Engle & Kane, 2004). According to such views, control may be accomplished proactively, in advance of interference or conflict, via goal-maintenance mechanisms that are vulnerable to disruption by external or internal distractors. Or, it may be accomplished reactively, in the face of experienced interference of conflict, via competition-resolution mechanisms that take real time to overcome the inertia of long- or short-term habit. Both of these mechanisms seem to vary with WMC (e.g., Kane & Engle, 2003), but if TUTs primarily reflect (or cause) goal-maintenance failures, then WMC's TUT-independent prediction of performance may represent the contributions of competition-resolution processes. Our vigilance SART tested this idea by requiring subjects to not respond to most trials, thus removing the go prepotency—and the importance of competition resolution—from the task. With goal maintenance being the primary control process left to affect performance, we predicted TUT rate to more fully mediate WMC's prediction of vigilance SART outcomes.

We were wrong. Instead, WMC predicted neither performance nor thought content in this task. Why? We cannot easily explain it via coordinative binding theory, because the standard and vigilance SARTs don't differ in their S–R binding rules. In both tasks, subjects must learn and maintain the identical, very simple, S–R rule (e.g., “if animal name, press key; if food name, don't press”). Where the standard and vigilance SART differ is in the momentary implementation of that rule, with standard SART being considerably more difficult, we claim, due to the prepotency to respond that is built over trials and that must be actively controlled for critical no-go stimuli. With both tasks involving the same S–R mapping, then, the binding view must make some additional assumptions in order to explain WMC's differential prediction of performance in each. For example, the no-go version of the rule might be somehow more difficult to bind (or learn) than is the go version, or the mapping is somehow more difficult to maintain over the course of the no-go task than the go task. Our study cannot rule out these claims, and so if they are empirically testable, they are worth investigating further. Nevertheless, we emphasize that coordinative binding theory has little to say about TUTs or attentional lapses, other than to claim that they may not be important to WMC variation (Schmiedek et al., 2007), and so it cannot help explain WMC's differential relation to TUTs across the two different SARTs.

The executive attention view may also explain the presence of a WMC effect in only standard SART but also, admittedly, in a post hoc fashion. We speculate that instead of merely reducing the demand for the competition-resolution component of executive attention, the vigilance SART also eliminated subjects' use of proactive executive control altogether. Subjects may have adopted a bottom-up attention strategy that allowed environmental events (i.e., target onsets) to trigger their reactions rather than allocating their attention proactively to maintain the task goal (Johnson et al., 2007), and so WMC no longer predicted off-task thinking or task accuracy. Braver et al. (2007) have pointed out significant disadvantages to proactive control that may encourage subjects to adopt a reactive stance during some tasks: Proactive control is resource-demanding (and metabolically costly; see also Gailliot et al., 2007), and it interferes with the automatization process. Based on these costs, Braver et al. argued, the cognitive system trades off between proactive and reactive control, and conflict cues must be quite strong and highly predictive in order to initiate costly proactive processes.

Indeed, null effects of WMC on attention-demanding tasks are not anomalous, and we have argued elsewhere that discovering such boundary conditions in the relation between WMC and “attention” is critical to inductively advancing the understanding of both WMC and executive control (Kane, Poole, Tuholski, & Engle, 2006; see also Barrouillet, Lépine, & Camos, 2008; Redick, Calvo, Gay, & Engle, 2011). Initially, individual-differences research seemed to suggest that WMC predicted performance in tasks generally thought to require controlled processing but not in those allowing automatic processing (e.g., Conway & Engle, 1994; Kane et al., 2001; Kane & Engle, 2000; Unsworth et al., 2004). Subsequent work demonstrated, however, that even quite difficult tasks requiring top-down control could be immune to WMC's influence, such as many varieties of visual search, including feature-absence, feature-conjunction, and command

searches (see Kane et al., 2006; Poole & Kane, 2009; Sobel, Gerrie, Poole, & Kane, 2007). Obviously, the vigilance SART is not a visual search task, but it has in common with these tasks the need to identify targets amid presentations of nontargets without the need to withhold prepotent responses or block distractor processing. Perhaps the restraint of habitual action or the constraint of conscious focus is critical to a task's eliciting WMC-related performance differences (Kane et al., 2006; but see Colflesh & Conway, 2007). Regarding thought content, at least one other study (a daily-life, experience-sampling study) has shown that lower WMC subjects mind-wander more than do higher WMC subjects when they report trying hard to concentrate on their ongoing activity but not when they report little effort to concentrate (Kane, Brown, et al., 2007). It is possible, then, that particular task features, such as the need to actively prevent commission errors, induce higher WMC subjects into proactive control modes that serve to combat off-task thinking (see also Smallwood, 2010). In any case, it is clear that future work on WMC and mind wandering should systematically manipulate a variety of tasks' executive demands in order to further clarify the boundaries of WMC- and TUT-related effects on cognition and cognitive individual differences.

WMC, attentional lapses, and the worst performance rule

Schmiedek et al. (2007) and Unsworth et al. (2010) both demonstrated the worst performance rule by connecting individual differences in WMC to the τ parameter from subjects' individualized RT distributions. They differed, however, in their interpretation of these results. Unsworth et al., like many investigators (e.g., Coyle, 2003; Larson & Alderton, 1990), attributed subjects' longest RTs to lapses of attention to ongoing task demands and thus argued that their findings supported the executive attention theory of WMC. Schmiedek et al., in contrast, argued that τ and its covariation with WMC reflected general information-processing efficiency, rather than an influence of attentional lapses (consistent with coordinative binding theory); moreover, they backed their claim via formal modeling (à la Ratcliff, Schmiedek, & McKoon, 2008; Wagenmakers et al., 2007) and a successful simulation study. We find no fault with the Schmiedek et al. methods or analyses, and their simulation results were compelling as far as they went. Our view, however, is that one can draw the strongest conclusions about attentional lapses by assessing them, rather than by inferring their unimportance via null modeling results (where the model includes no parameter that corresponds to lapses).

When we actually measured attentional lapses via probed thought reports and the drift-rate variability parameter from the LBA model, we found that TUT rate and drift-rate variability significantly predicted normal variation in τ . Of most importance here, both TUT rate and drift-rate variability also partially mediated the correlation between WMC and τ . That is, individual differences in the subjective experience of attentional lapses, and in the intertrial variability in evidence accumulation left in the wake of those lapses, predicted long RTs in a continuous go/no-go task. Variation in attentional lapses was, moreover, partly responsible for WMC's prediction of long RTs. We should make clear that neither result falsifies the Schmiedek et al. (2007) claim that individual differences in τ are driven to some considerable extent by

differences in general information-processing efficiency that might reflect S–R binding capability, at least in some task contexts (see also Ratcliff et al., 2008). Our results, however, directly support the claims of executive attention theory (e.g., Kane, Conway, et al., 2007) that WMC's predictive power derives—to some degree—from its tapping into variation in attention-control processes involved in the regulation of both thought and behavior.

We must note, however, that our modeling results differ from previously published reports on individual differences in RT tasks. In line with Schmiedek et al. (2007), this small but growing literature converges on the idea that drift rate is especially important to ability-related individual differences in choice-task performance. First, simulation studies (Ratcliff et al., 2008; van Ravenzwaaij, Brown, & Wagenmakers, 2011) have linked drift rate to the worst performance rule and to the association between intelligence and RT (and RT variability). Second, empirical work has shown that IQ (as assessed by psychometric tests of matrix reasoning and vocabulary) correlates significantly and strongly with drift rate, especially in younger adults (Ratcliff et al., 2010; Ratcliff, Thapar, & McKoon, 2011), with weak to null correlations with response criterion and nondecision parameters. What might account for our discrepant findings? We consider two possibilities as most promising.

First, most of the previous studies have considered only drift rate, response criterion, and nondecision times to be the parameters of interest and have not sought to model or test for any role for drift-rate variability in producing ability-related individual differences (although drift-rate variability is needed by the diffusion model to account for the RT distributions of errors relative to accurate responses, it and the other within-subject variability parameters are often not accorded psychological or process-based interpretations). Second, the SART seems to differ from most choice-RT tasks that have served as the basis for diffusion modeling in individual-differences studies. It is a highly repetitive task that offers subjects no breaks over the course of a 45-min session, and it encourages rapid and mindless responding that seems to elicit both very fast and very slow responses (and so highly variable responding). Given that evidence-accumulation modeling has been successfully employed with go/no-go tasks (Gomez et al., 2007) and that LBA modeling yields parameter estimates similar to those for diffusion modeling when they are applied to the same data (Donkin et al., 2011), we do not believe that our findings are idiosyncratic to our modeling methods. Future work will be necessary, however, to further examine the contributions of drift-rate variability to ability-related individual differences across different varieties of tasks and subject groups.

WMC and theoretical accounts of mind wandering

A current debate in mind-wandering research concerns the role of executive processes, or resources. The Smallwood and Schooler (2006) view, along with numerous empirical reports (e.g., Ellis, Moore, Varner, Ottaway, & Becker, 1997; Forster & Lavie, 2009; Giambra, 1989; McKiernan et al., 2006; Teasdale et al., 1995), has argued that TUTs consume executive resources. McVay and Kane (2010) countered these claims with a Control Failure × Concerns

view, based largely on individual differences in the propensity to mind-wander and their connection to executive control. The Control Failure \times Concerns view posits that unwanted TUTs during ongoing tasks reflect a failure to control attention and maintain task goals in the face of interference from task-irrelevant, concern-related thoughts that are automatically cued by environmental or mental events. An important hypothesis derived from this perspective is that people with deficient control capabilities will more often succumb to TUTs than will those with better control (as will those who have more vs. less urgent personal concerns with which to contend). The resource-demanding view of mind wandering (e.g., Giambra, 1989; Smallwood & Schooler, 2006), in contrast, makes the opposite prediction. Namely, people with more executive resources at their disposal should mind-wander more than those with fewer. That is, if mind wandering is resource-demanding, and if ongoing tasks and TUTs draw on the same resource pool, then people with greater resources should more effectively balance TUTs and on-task thinking (and thus, performance).

The current study provides evidence for the Control Failure \times Concerns view of mind wandering (McVay & Kane, 2010). First, WMC correlated negatively with TUT rate during an attention-demanding task (see also Kane, Brown, et al., 2007; McVay & Kane, 2009, 2011). Second, if mind wandering were resource-demanding, then higher WMC subjects' performance should be affected to a lesser degree than lower WMC subjects', as higher WMC provides more resources to distribute between task performance and TUTs. A reexamination of data from McVay and Kane (2009; reported in McVay & Kane, 2010) found that, overall, subjects were less accurate when mind wandering but that the task performance of high and low WMC subjects were affected to the same degree by TUTs. The current study replicated this finding: Although subjects were more likely to make a standard SART error when they reported a TUT, WMC did not interact significantly with thought report, signifying that high and low WMC subjects (defined as top and bottom quartile scorers) experienced the same performance decrement when mind wandering ($M_s = .83$ vs. $.59$, respectively, for high WMC and $M_s = .74$ vs. $.57$, respectively, for low WMC). 6

Although our findings seem inconsistent with prototypical resource views of mind wandering, Smallwood's (2010) response to McVay and Kane (2010) provided an alternative conception of executive "resources" and their consumption during mind-wandering states. TUTs, here, occupy the global workspace of consciousness (e.g., Baars, 1988; Navon, 1989a, 1989b). According to workspace theories, modular processing networks that are specialized for particular functions can be brought under general executive control when goals or other representations are made globally available to the cognitive system via consciousness (i.e., reportable experiences). Smallwood's logic, then, is that (1) access to global broadcasting is capacity limited and (2) TUTs, as conscious experiences, occupy the workspace, ergo (3) TUTs must consume an executive resource. Moreover, Smallwood argued that this workspace view correctly predicts the all-or-none pattern we found that experiencing a TUT in the moment should impair performance similarly regardless of a subject's executive control abilities.

On one hand, then, Smallwood's (2010) argument jibes with our view that conscious access to goals is critical to the executive regulation of behavior and thought and that automatically cued thoughts about personal concerns may commandeer consciousness and thus thwart attempts at proactive and reactive control. From this perspective, TUTs do appear to influence, if not engage, executive mechanisms. On the other hand, we still differ fundamentally from Smallwood in our claim that executive control mechanisms are primarily important in preventing TUTs from gaining access to the global workspace in the first place; that is, TUTs can monopolize the workspace-as-executive-mechanism only if they are allowed to intrude into consciousness by inefficient control processes. Moreover, executive control processes are heterogeneous and encompass much more than just the global workspace (e.g., Botvinick et al., 2001; Braver et al., 2007; Kane & Engle, 2003; Miller & Cohen, 2001; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000), and so it would be unwise to characterize TUTs as consuming all or most of some undifferentiated executive "resource" (for more general concerns about the viability of resource views of attention, see Navon, 1984; Neumann, 1987).

Objective Markers, and Validity, of Subjective Mind-Wandering Reports

The search for a reliable, objective marker of mind wandering is ongoing. Unfortunately, the dynamic patterns of changing RTs originally identified by Smallwood, McSpadden, et al. (2008) did not reliably predict the occurrence of TUTs in our large data set that combined our new data with those from McVay and Kane (2009). Using long pretarget RT series, all standardized within-subject, we replicated their finding of three principal components. Of importance, all three of the identified RT patterns predicted target no-go errors (i.e., series that were generally faster than average and those that changed from relatively slow to relatively fast); we therefore replicated the Smallwood group's interesting finding that impending errors can be predicted well in advance by RT changes. That same within-series change, from slow to fast, did not reliably predict TUTs, however. This difference in RTs' predictions of impending errors versus thought reports further supports our general argument that errors are not always isomorphic with TUTs and that errors are determined only in part (or only some of the time) by lapses of thought. In short, theoretical inferences about thought content cannot reliably be made based on performance patterns alone (see also Helton et al., 2005, 2010). It seems that further research will be necessary to establish objective behavioral markers of mind wandering (but for other promising behavioral, psychophysiological, and neuroimaging findings, see Christoff et al., 2009; Reiche, Reineberg, & Schooler, 2010; Smallwood, Beach, et al., 2008; Smallwood et al., 2004; Smallwood, O'Connor, et al., 2007). In the meantime, the thought probe remains a useful, if subjective, tool for testing the importance of thought content in theoretical discussions of WMC and executive control.

Indeed, probed reports of TUT experiences have a strong record of validity, as we already reviewed. The present study provides additional sources of validation, including evidence that TUT reports did not simply reflect subjects' reactive, post hoc explanations for their own performance. For example, if subjects had used target-trial performance to determine whether

their mind had wandered (“Oops, I missed that one. I must have been mind-wandering”), then target accuracy should have predicted thought reports more strongly than it did. Although performance varied systematically with thought reports, subjects committed errors on 38% of trials preceding on-task thought reports (in the standard SART; similar to the 34% reported in McVay & Kane, 2009) and responded accurately on 36% of trials preceding TUT reports (42% in McVay & Kane, 2009); clearly, subjects often reported thoughts at odds with their performance. Similarly, regarding individual differences, the significant but nonperfect correlation between TUT rate and dL indicates that most of the variance in thought reports was independent of SART accuracy. Furthermore, in both McVay and Kane (2009) and the current study, overall TUT rate correlated as strongly with intraindividual RT variation as with SART accuracy, but only the latter could be easily monitored and used to influence thought reports. Indeed, our quantitative diffusion-modeling results also refute the reactivity of TUT reports, because the drift-rate-variability parameter from the LBA model correlated with WMC and TUT reports, and it partially mediated the WMC–TUT and WMC– τ associations: Subjects obviously could not use their drift-rate-variability parameter to inform their thought reports!

Finally, key differences between subjects' reports of TUTs and TRI (i.e., evaluative thoughts about their performance) offer further evidence for thought-report validity. Although McVay and Kane (2009) did not report analyses of TRI, in reexamining those data here we note that standard SART accuracy was nearly as low for trials on which subjects reported TRI ($M = .44$) as it was for those on which they reported TUTs ($M = .42$), suggesting that both varieties of off-task thought hurt performance (vs. $M = .66$ for on-task trials). In the current study, the standard SART patterns for accuracy were similar (TRI $M = .38$; TUT $M = .36$; on-task $M = .62$). At the same time, the latency with which subjects indicated their thought content to probes was different for TUTs and TRI. Subjects more quickly categorized their thoughts as being on-task or about their task performance ($M_s \pm SEMs = 2,309 \pm 57$ ms vs. $2,705 \pm 67$ ms for on-task vs. TRI reports) than as being about task-unrelated topics (TUTs: $3,268 \pm 80$ ms). A reanalysis of the McVay and Kane (2009) data shows a similar pattern ($M_s \pm SEMs$ for on-task thoughts = $2,213 \pm 60$ ms, for TRI = $2,261 \pm 48$ ms, and for TUTs = $2,665 \pm 59$ ms). If subjects simply allowed their accuracy to influence their thought reports, they should have made both TUT and TRI responses (indicative of error) with similar ease and speed.

Conclusions

The measurement of mind wandering, or TUTs, within a task contributes significantly to the understanding of individual differences in WMC and attention control. The negative correlation between WMC and TUT rate supports the executive attention theory of WMC, which claims that a primary factor underlying both tests of WMC and complex cognition (e.g., reading comprehension, scholastic achievement tests, and fluid intelligence tests) is executive control. Furthermore, our thought-report and evidence-accumulation modeling findings indicate that lapses of attention contribute to the worst performance rule, whereby subjects' longest RTs (and the ex-Gaussian τ parameter) correlate most strongly with cognitive ability. More broadly, the

apparent impact of off-task thoughts on particular varieties of task performance demands a closer look at the ways in which thought control and action control interact to produce goal-directed behavior (see e.g., the hypothesized addition of a Supervisory Attention Gateway to classic models of the Supervisory Attention System; Burgess, Dumontheil, & Gilbert, 2007; Gilbert, Frith, & Burgess, 2005; Gilbert, Simons, Frith, & Burgess, 2006).

Footnotes

1 We have also proposed that control may be implemented reactively, by mechanisms specialized for resolving response conflict and distractor inhibition (Engle & Kane, 2004; Kane, Conway, et al., 2007; Kane & Engle, 2003; see also Braver et al., 2007; Hasher, Lustig, & Zacks, 2007; Jacoby et al., 1999).

2 The diffusion model parameters of most consistent interest across studies are evidence accumulation/drift rate (ν), response criterion/boundary separation (a), nondecision components (T_{er}), drift starting point/response bias (z), across-trial variation in drift rate (η), across-trial variation in nondecision components (σ_{st}), across-trial variation in starting point (σ_{sz}), and, most recently, proportion of contaminant RTs (p_o ; e.g., Ratcliff & Tuerlinckx, 2002).

3 Even if binding theory (Oberauer et al., 2007; Wilhelm & Oberauer, 2006) were to claim that the standard SART challenges S–R binding capabilities more than does the vigilance SART (see our General Discussion for more on this issue), the vigilance SART variance that is predicted by WMC should still have no association with TUT rate.

4 The 12 trials used for analysis in Smallwood, McSpadden, et al. (2008) differed between their slow and fast pace conditions. All 12 discrete trials were used in the slow condition, whereas 12 “averaged” RTs were used in the fast; the fast condition presented 24 trials per block, so the authors averaged across consecutive trials to equate the number of trials between pace conditions.

5 McVay and Kane (2009) found that TRI comprised 24% of the standard SART thought-probe responses, but they did not analyze TRI reports further. Analyzing those data here yielded a similar decrease in TRI rate over blocks, a similar prediction of in-the-moment error versus on-task thought reports, and similar null correlations between individual differences in TRI rate and SART performance. The only discrepancy with the present findings was a significant WMC \times TRI correlation ($r = .19$), such that higher WMC subjects reported higher rates of TRI than did lower WMC subjects. Given the failure to replicate this positive correlation here, we do not consider it further.

6 It is also possible that higher and lower WMC subjects are equivalently hurt by TUTs because higher WMC subjects entertain more complex (i.e., more resource-demanding) thoughts during tasks than do lower WMC subjects (J. Smallwood, personal communication, October 2010). It would be difficult to test such a claim, however, and we are skeptical that higher WMC subjects

would engage in thought just complex enough to make their “dual-tasking” cost similar to lower WMC subjects'. It seems to us much more likely that engaging in off-task thought of any kind induces a cost to ongoing performance that is of similar magnitude for everyone (see also Smallwood, 2010).

References

- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, 80, 606–621. doi: 10.1111/j.1467-8624.2009.01282.x
- Antrobus, J. S., Coleman, R., & Singer, J. L. (1967). Signal-detection performance by subjects differing in predisposition to daydreaming. *Journal of Consulting Psychology*, 31, 487– 491. doi: 10.1037/h0024969
- Antrobus, J. S., Singer, J. L., & Greenberg, S. (1966). Studies in the stream of consciousness: Experimental enhancement and suppression of spontaneous cognitive processes. *Perceptual and Motor Skills*, 23, 399– 417. doi: 10.2466/pms.1966.23.2.399
- Baars, B. J. (1988). *A cognitive theory of consciousness*. New York, NY: Cambridge University Press.
- Barrouillet, P., L epine, R., & Camos, V. (2008). Is the influence of working memory capacity on high-level cognition mediated by complexity or resource-dependent elementary processes? *Psychonomic Bulletin & Review*, 15, 528– 534. doi: 10.3758/PBR.15.3.528
- Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, 130, 701– 725. doi: 10.1037/0096-3445.130.4.701
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. C. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624– 652. doi: 10.1037/0033-295X.108.3.624
- Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 76– 106). New York, NY: Oxford University Press.
- Brown, S., & Heathcote, A. (2003). QMLE: Fast, robust, and efficient estimation of distribution functions based on quantiles. *Behavior Research Methods, Instruments & Computers*, 35, 485– 492. doi: 10.3758/BF03195527
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57, 153– 178.

- Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Sciences*, 11, 290– 298. doi: 10.1016/j.tics.2007.05.004
- Cheyne, J. A., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, 111, 98– 113. doi: 10.1016/j.cognition.2008.12.009
- Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences, USA*, 106, 8719– 8724. doi: 10.1073/pnas.0900234106
- Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review*, 99, 45– 77. doi: 10.1037/0033-295X.99.1.45
- Colflesh, G. J. H., & Conway, A. R. A. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic Bulletin & Review*, 14, 699– 703. doi: 10.3758/BF03196824
- Colom, R., Abad, F. J., Quiroga, M. A., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence*, 36, 584– 606. doi: 10.1016/j.intell.2008.01.002
- Colzato, L. S., Spapé, M. M. A., Pannebakker, M. M., & Hommel, B. (2007). Working memory and the attentional blink: Blink size is predicted by individual differences in operation span. *Psychonomic Bulletin & Review*, 14, 1051– 1057. doi: 10.3758/BF03193090
- Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review*, 8, 331– 335. doi: 10.3758/BF03196169
- Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, 123, 354– 373. doi: 10.1037/0096-3445.123.4.354
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12, 769– 786. doi: 10.3758/BF03196772
- Coyle, T. R. (2003). A review of the worst performance rule: Evidence, theory, and alternative hypotheses. *Intelligence*, 31, 567– 587. doi: 10.1016/S0160-2896(03)00054-0

- Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3, 422– 433. doi: 10.3758/BF03214546
- Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. New York, NY: Academic Press.
- De Jong, R. D. (2001). Adult age differences in goal activation and goal maintenance. *European Journal of Cognitive Psychology*, 13, 71– 89. doi: 10.1080/09541440042000223
- De Jong, R. D., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica*, 101, 379– 394. doi: 10.1016/S0001-6918(99)00012-8
- Donkin, C., Averell, L., Brown, S., & Heathcote, A. (2009). Getting more from accuracy and response time data: Methods for fitting the linear ballistic accumulator. *Behavior Research Methods*, 41, 1095– 1110. doi: 10.3758/BRM.41.4.1095
- Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E.-J. (2011). Diffusion versus linear ballistic accumulation: Different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review*, 18, 61– 69. doi: 10.3758/s13423-010-0022-4
- Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. Gazzaniga (Ed.) , *The cognitive neurosciences* (pp. 721– 733). Cambridge, MA: MIT Press.
- Ellis, H. C., Moore, B. A., Varner, L. J., Ottaway, S. A., & Becker, A. S. (1997). Depressed mood, task organization, cognitive interference, and memory: Irrelevant thoughts predict recall performance. *Journal of Social Behavior & Personality*, 12, 453– 470.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.) , *The psychology of learning and motivation* (pp. 145– 199). New York, NY: Academic Press.
- Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition*, 111, 345– 355. doi: 10.1016/j.cognition.2009.02.006
- Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *Journal of Neuroscience*, 29, 8726– 8733. doi: 10.1523/JNEUROSCI.2145-09.2009
- Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., . . . Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, 92, 325– 336. doi: 10.1037/0022-3514.92.2.325

- Gathercole, S. E., Alloway, T. P., Kirkwood, H. J., Elliott, J. G., Holmes, J., & Hilton, K. A. (2008). Attentional and executive function behaviours in children with poor working memory. *Learning and Individual Differences*, 18, 214– 223. doi: 10.1016/j.lindif.2007.10.003
- Giambra, L. M. (1989). Task-unrelated thought frequency as a function of age: A laboratory study. *Psychology and Aging*, 4, 136– 143. doi: 10.1037/0882-7974.4.2.136
- Giambra, L. M. (1993). The influence of aging on spontaneous shifts of attention from external stimuli to the contents of consciousness. *Experimental Gerontology*, 28, 485– 492. doi: 10.1016/0531-5565(93)90073-M
- Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition*, 4, 1– 21. doi: 10.1006/ccog.1995.1001
- Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2005). Involvement of rostral prefrontal cortex in selection between stimulus-oriented and stimulus-independent thought. *European Journal of Neuroscience*, 21, 1423– 1431. doi: 10.1111/j.1460-9568.2005.03981.x
- Gilbert, S. J., Simons, J. S., Frith, C. D., & Burgess, P. W. (2006). Performance-related activity in medial rostral prefrontal cortex (area 10) during low-demand tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 45– 58. doi: 10.1037/0096-1523.32.1.45
- Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, 136, 389– 413.
- Grodsky, A., & Giambra, L. M. (1990–1991). The consistency across vigilance and reading tasks of individual differences in the occurrence of task-unrelated and task-related images and thoughts. *Imagination, Cognition and Personality*, 10, 39– 52.
- Hasher, L., Lustig, C., & Zacks, R. T. (2007). Inhibitory mechanisms and the control of attention. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 227– 249). Oxford, England: Oxford University Press.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193– 225). San Diego, CA: Academic Press.
- Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General*, 136, 217– 240. doi: 10.1037/0096-3445.136.2.217
- Helton, W. S., Hollander, T. D., Warm, J. S., Matthews, G., Dember, W. N., Wallaart, M., . . . Hancock, P. A. (2005). Signal regularity and the mindlessness model of vigilance. *British Journal of Psychology*, 96, 249– 261. doi: 10.1348/000712605X38369

Helton, W. S., Weil, L., Middlemiss, A., & Sawers, A. (2010). Global interference and spatial uncertainty in the sustained attention to response task (SART). *Consciousness and Cognition*, 19, 77– 85. doi: 10.1016/j.concog.2010.01.006

Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 851– 860.

Jackson, J. D., & Balota, D. A. (2011, June27). Mind-wandering in younger and older adults: Converging evidence from the sustained attention to response task and reading for comprehension. *Psychology and Aging*. Advance online publication. doi: 10.1037/a0023933

Jacoby, L. L., Kelley, C. M., & McElree, B. D. (1999). The role of cognitive control: Early selection versus late correction. In S. Chaiken & E. Trope (Eds.) , *Dual process theories in social psychology* (pp. 383– 400). New York, NY: Guilford Press.

Jensen, A. R. (1992). The importance of intraindividual variability in reaction time. *Personality and Individual Differences*, 13, 869– 881. doi: 10.1016/0191-8869(92)90004-9

Johnson, K. A., Kelly, S. P., Bellgrove, M. A., Barry, E., Cox, M., Gill, M., . . . Robertson, I. H. (2007). Response variability in attention deficit hyperactivity disorder: Evidence for neuropsychological heterogeneity. *Neuropsychologia*, 45, 630– 638. doi: 10.1016/j.neuropsychologia.2006.03.034

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, 130, 169– 183. doi: 10.1037/0096-3445.130.2.169

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science*, 18, 614– 621. doi: 10.1111/j.1467-9280.2007.01948.x

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.) , *Variation in working memory* (pp. 21– 48). New York, NY: Oxford University Press.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 336– 358. doi: 10.1037/0278-7393.26.2.336

- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132, 47–70. doi: 10.1037/0096-3445.132.1.47
- Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131, 66–71. doi: 10.1037/0033-2909.131.1.66
- Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of “executive attention.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 749–777. doi: 10.1037/0278-7393.32.4.749
- Keye, D., Wilhelm, O., Oberauer, K., & van Ravenzwaaij, D. (2009). Individual differences in conflict-monitoring: Testing means and covariance hypothesis about the Simon and the Eriksen Flanker task. *Psychological Research*, 73, 762–776. doi: 10.1007/s00426-008-0188-9
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.
- Klinger, E., & Cox, W. M. (1987–1988). Dimensions of thought flow in everyday life. *Imagination, Cognition and Personality*, 7, 105–128.
- Larson, G. E., & Alderton, D. L. (1990). Reaction time variability and intelligence: A “worst performance” analysis of individual differences. *Intelligence*, 14, 309–325. doi: 10.1016/0160-2896(90)90021-K
- Larson, G. E., & Saccuzzo, D. P. (1989). Cognitive correlates of general intelligence: Toward a process theory of g. *Intelligence*, 13, 5–31. doi: 10.1016/0160-2896(89)90003-2
- Long, D. L., & Prat, C. S. (2002). Working memory and Stroop interference: An individual differences investigation. *Memory & Cognition*, 30, 294–301. doi: 10.3758/BF03195290
- Marcovitch, S., Boseovski, J. J., Knapp, R. J., & Kane, M. J. (2010). Goal neglect and working memory capacity in 4- to 6-year-old children. *Child Development*, 81, 1687–1695. doi: 10.1111/j.1467-8624.2010.01503.x
- Martínez, K., & Colom, R. (2009). Working memory capacity and processing efficiency predict fluid but not crystallized and spatial intelligence: Evidence supporting the neural noise hypothesis. *Personality and Individual Differences*, 46, 281–286. doi: 10.1016/j.paid.2008.10.012
- Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007, April 25). Wandering minds: The default network and stimulus-independent thought. *Science*, 315, 393–395. doi: 10.1126/science.1131295

- Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, 16, 798–817. doi: 10.3758/PBR.16.5.798
- McGuire, P. K., Paulesu, E., Frackowiak, R. S. J., & Frith, C. D. (1996). Brain activity during stimulus independent thought. *NeuroReport*, 7, 2095–2099.
- McKiernan, K. A., D'Angelo, B. R., Kaufman, J. N., & Binder, J. R. (2006). Interrupting the stream of consciousness: An fMRI investigation. *NeuroImage*, 29, 1185–1191. doi: 10.1016/j.neuroimage.2005.09.030
- McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 196–204. doi: 10.1037/a0014104
- McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2005). and Watkins (2008). *Psychological Bulletin*, 136, 188–197. doi: 10.1037/a0018298
- McVay, J. C., & Kane, M. J. (2011). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*. Advance online publication. doi: 10.1037/a0025250
- McVay, J. C., Kane, M. J., & Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind-wandering across controlled and ecological contexts. *Psychonomic Bulletin & Review*, 16, 857–863. doi: 10.3758/PBR.16.5.857
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49–100. doi: 10.1006/cogp.1999.0734
- Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationship among working memory, secondary memory, and fluid intelligence. *Psychological Science*, 19, 1071–1077. doi: 10.1111/j.1467-9280.2008.02202.x
- Navon, D. (1984). Resources—A theoretical soup stone? *Psychological Review*, 91, 216–234. doi: 10.1037/0033-295X.91.2.216

Navon, D. (1989a). The importance of being visible: On the role of attention in a mind viewed as an anarchic intelligence system. I. Basic tenets. *European Journal of Cognitive Psychology*, 1, 191– 213.

Navon, D. (1989b). The importance of being visible: On the role of attention in a mind viewed as an anarchic intelligence system. II. Application to the field of attention. *European Journal of Cognitive Psychology*, 1, 215– 238.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H.Heuer & A. F.Sanders (Eds.) , *Perspectives on perception and action* (pp. 361– 394). Hillsdale, NJ: Erlbaum.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231– 259. doi: 10.1037/0033-295X.84.3.231

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, 134, 368– 387. doi: 10.1037/0096-3445.134.3.368

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, 51, 45– 100. doi: 10.1016/S0079-7421(09)51002-X

Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H. M. (2005). Working memory and intelligence—their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, 131, 61– 65. doi: 10.1037/0033-2909.131.1.61

Oberauer, K., Süß, H. M., Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. In A. R. A.Conway, C.Jarrold, M. J.Kane, A.Miyake, & J. N.Towse (Eds.) , *Variation in working memory* (pp. 49– 75). New York, NY: Oxford University Press.

Perry, A. R., & Laurie, C. A. (1992). Sustained attention and the Type A behavior pattern: The effect of daydreaming on performance. *Journal of General Psychology*, 119, 217– 228. doi: 10.1080/00221309.1992.9917803

Poole, B. J., & Kane, M. J. (2009). Working memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *Quarterly Journal of Experimental Psychology*, 62, 1430– 1454. doi: 10.1080/17470210802479329

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9, 347– 356. doi: 10.1111/1467-9280.00067

Ratcliff, R., Schmiedek, F., & McKoon, G. (2008). A diffusion model explanation of the worst performance rule for reaction time and IQ. *Intelligence*, 36, 10– 17. doi: 10.1016/j.intell.2006.12.002

- Ratcliff, R., & Smith, P. L. (2004). A comparison of serial sampling models for two-choice reaction time. *Psychological Review*, 111, 333– 367. doi: 10.1037/0033-295X.111.2.333
- Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology*, 60, 127– 157. doi: 10.1016/j.cogpsych.2009.09.001
- Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General*, 140, 464– 487. doi: 10.1037/a0023810
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9, 438– 481. doi: 10.3758/BF03196302
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.
- Reason, J. T. (1990). *Human error*. Cambridge, England: Cambridge University Press.
- Redick, T. S., Calvo, A., Gay, C. E., & Engle, R. W. (2011). Working memory capacity and go/no-go task performance: Selective effects of updating, maintenance, and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 308– 324. doi: 10.1037/a0022216
- Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science*, 21, 1300– 1310.
- Roberts, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology*, 12, 105– 126. doi: 10.1080/87565649609540642
- Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). “Oops!”: Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, 35, 747– 758. doi: 10.1016/S0028-3932(97)00015-8
- Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H. M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, 136, 414– 429. doi: 10.1037/0096-3445.136.3.414
- Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences*, 6, 339– 344. doi: 10.1016/S1364-6613(02)01949-6

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Zoning out while reading: Evidence for dissociations between experience and metaconsciousness. In D. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 203–226). Cambridge, MA: MIT Press.

Shaw, G. A., & Giambra, L. M. (1993). Task unrelated thoughts of college students diagnosed as hyperactive in childhood. *Developmental Neuropsychology*, 9, 17–30. doi: 10.1080/87565649309540541

Smallwood, J. (2010). Why the global availability of mind wandering necessitates resource competition: Reply to McVay & Kane. *Psychological Bulletin*, 136, 202–207.

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience*, 20, 458–469. doi: 10.1162/jocn.2008.20037

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition*, 13, 657–690. doi: 10.1016/j.concog.2004.06.003

Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009). Shifting moods, wandering minds: Negative moods lead the mind to wander. *Emotion*, 9, 271–276. doi: 10.1037/a0014855

Smallwood, J., McSpadden, M., Luus, B., & Schooler, J. (2008). Segmenting the stream of consciousness: The psychological correlates of temporal structures in the time series data of a continuous performance task. *Brain and Cognition*, 66, 50–56. doi: 10.1016/j.bandc.2007.05.004

Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review*, 14, 527–533. doi: 10.3758/BF03194102

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incidence of the wandering mind. *Memory & Cognition*, 36, 1144–1150. doi: 10.3758/MC.36.6.1144

Smallwood, J., Obonsawin, M., Baracaia, S. F., Reid, H., O'Connor, R., & Heim, D. (2002–2003). The relationship between rumination, dysphoria, and self-referent thinking: Some preliminary findings. *Imagination, Cognition, and Personality*, 22, 317–342. doi: 10.2190/2N80-AVM3-4A23-LEAJ

Smallwood, J., Obonsawin, M., & Reid, H. (2003). The effects of block duration and task demands on the experience of task unrelated thought. *Imagination, Cognition and Personality*, 22, 13–31. doi: 10.2190/TBML-N8JN-W5YB-4L9R

Smallwood, J., O'Connor, R. C., Sudbery, M. V., & Obonsawin, M. C. (2007). Mind-wandering and dysphoria. *Cognition & Emotion*, 21, 816– 842. doi: 10.1080/02699930600911531

Smallwood, J., Riby, L., Heim, D., & Davies, J. D. (2006). Encoding during the attentional lapse: Accuracy of encoding during the semantic SART. *Consciousness and Cognition*, 15, 218– 231. doi: 10.1016/j.concog.2005.03.003

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, 132, 946– 958. doi: 10.1037/0033-2909.132.6.946

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34– 50. doi: 10.1037/0096-3445.117.1.34

Sobel, K. V., Gerrie, M. P., Poole, B. J., & Kane, M. J. (2007). Individual differences in working memory capacity and visual search: The roles of top-down and bottom-up processing. *Psychonomic Bulletin & Review*, 14, 840– 845. doi: 10.3758/BF03194109

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245– 251. doi: 10.1037/0033-2909.87.2.245

Stuyven, E., & Van der Goten, K. (1995). Stimulus independent thoughts and working memory: The role of the central executive. *Psychologica Belgica*, 35, 241– 251.

Teasdale, J. D., Dritschel, B. H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., & Baddeley, A. D. (1995). Stimulus-independent thought depends on central executive resources. *Memory & Cognition*, 23, 551– 559. doi: 10.3758/BF03197257

Teasdale, J. D., Proctor, L., Lloyd, C. A., & Baddeley, A. D. (1993). Working memory and stimulus-independent thought: Effects of memory load and presentation rate. *European Journal of Cognitive Psychology*, 5, 417– 433. doi: 10.1080/09541449308520128

Tse, C.-S., Balota, D. A., Yap, M. J., Duchek, J. M., & McCabe, D. P. (2010). Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology*, 24, 300– 315. doi: 10.1037/a0018274

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, 114, 104– 132. doi: 10.1037/0033-295X.114.1.104

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37, 498– 505. doi: 10.3758/BF03192720

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the

relationship between processing and storage. *Memory*, 17, 635– 654. doi: 10.1080/09658210902998047

Unsworth, N., Redick, T. S., Lakey, C. E., & Young, D. L. (2010). Lapses in sustained attention and their relation to executive control and fluid abilities: An individual differences investigation. *Intelligence*, 38, 111– 122. doi: 10.1016/j.intell.2009.08.002

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 1302– 1321. doi: 10.1037/0278-7393.30.6.1302

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*, 62, 392– 406. doi: 10.1016/j.jml.2010.02.001

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*, 14, 1011– 1026.

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, 40, 61– 72.

van Ravenzwaaij, D., Brown, S., & Wagenmakers, E. J. (2011). An integrated perspective on the relation between response speed and intelligence. *Cognition*, 119, 381– 393. doi: 10.1016/j.cognition.2011.02.002

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, 39, 767– 775.

Wagenmakers, E. J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, 14, 3– 22. doi: 10.3758/BF03194023

West, R. (2001). The transient nature of executive control processes in younger and older adults. *European Journal of Cognitive Psychology*, 13, 91– 105. doi: 10.1080/09541440042000232

Wilhelm, O., & Oberauer, K. (2006). Why are reasoning ability and working memory capacity related to mental speed? An investigation of stimulus–response compatibility in choice reaction time tasks. *European Journal of Cognitive Psychology*, 18, 18– 50. doi: 10.1080/09541440500215921

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Harvard University Press.

