American Psychological Association

Journal of Experimental Psychology: Applied

ISSN: 1076-898X

https://doi.org/10.1037/xap0000388

The Persistence of Distraction: The Hidden Costs of Intermittent Multitasking

Q1 David L. Strayer[1], Spencer C. Castro[2], Jonna Turrill[3], and Joel M. Cooper[4]
Q2 [1] Department of Psychology, University of Utah
Q3 [2] University of California, Merced
Q4 [3] Qualtrics
[4] Red Scientific, Inc.

We examined the hidden costs of intermittent multitasking. Participants performed a pursuit-tracking task (Experiment 1) or drove in a high-fidelity driving simulator (Experiment 2) by itself or while concurrently performing an easy or difficult backwards counting task that periodically started and stopped, creating on-task and off-task multitasking epochs. A novel application of the Detection Response Task (DRT), a standardized protocol for measuring cognitive workload (ISO 17488, 2016), was used to measure performance in the on-task and off-task intervals. We found striking costs that persisted well after the counting task had stopped. In fact, the multitasking costs dissipated as a negatively accelerated function of time with the largest costs observed immediately after multitasking ceased. Performance in the off-task interval remained above baseline levels throughout the 30-s off-task interval. We suggest that loading new procedures into working memory occurs fairly quickly, whereas purging this information from working memory takes considerably longer.

Public Significance Statement

Driver distraction caused by performing an unrelated secondary task is a significant cause of motor vehicle crashes on the roadway. This research documents that the effects of multitasking last well after secondary-task interactions have finished.

Q5 *Keywords:* attention, multitasking, dual-task processing, driving

Multitasking is a common activity engaged in by the vast majority of the individuals in the industrialized world. One ubiquitous form of multitasking involves motorist's concurrent use of a smartphone to talk or text. In fact, at any given daylight moment, 9.7% of the driving public in the United States can be seen holding or manipulating their smartphone while operating a motor vehicle (Department of Transportation [DOT], 2019). Newer vehicles equipped with Bluetooth capability pairing the driver's smartphone to a vehicle's infotainment system make the number of multitasking motorists on the roadway considerably higher (Strayer et al., 2019). This pervasive activity distracts drivers and leads to increases in injuries and fatalities on our roadways (World Health Organization [WHO], 2011).

Most people are poor at multitasking, suffering performance decrements to one or more of the concurrently performed tasks, even though they may think they are good at it (Sanbonmatsu et al., 2013; Strayer et al., 2011; Watson & Strayer, 2010). To date, most research has focused on the concurrent aspects of dual-task performance with much less attention given to performance following a multitasking episode (i.e., when people stop multitasking and return to the single task). In this article, we report the use of a novel method for examining the postmultitasking interval to identify persistent costs associated with multitasking.

Multitasking and Dual-Task Performance

Multitasking in the context of driving involves the performance of two or more *functionally independent tasks* with separate goals, stimuli, and responses.[1] We contrast this definition with performing *interdependent tasks* that serve common or overlapping goals. Note that the driving task and the smartphone tasks are functionally independent; they have separate goals (transportation vs. communication) with distinct stimuli (the driving environment vs. the content of the conversation) and separate responses (e.g., steering and braking vs. talking and listening). Regan et al. (2011, p. 1776) defined the distraction created by performing functionally independent tasks (i.e., "driver diverted attention") as "the diversion of attention away from activities critical for safe driving toward a competing activity, which may result in insufficient or no attention

David L. Strayer https://orcid.org/0000-0002-1653-730X
Spencer C. Castro https://orcid.org/0000-0003-1394-0184
Jonna Turrill https://orcid.org/0000-0003-1741-0797
Correspondence concerning this article should be addressed to David L. Strayer, Department of Psychology, University of Utah, 380 S. 1530 E. Room 502, Salt Lake City, UT 84112, United States. Email: David.Strayer@utah.edu

[1] We use the terms multitask and dual-task interchangeably in this article. A dual-tasking situation is one in which participants perform two functionally independent tasks at the same time. Multitasking is a more generalized term in which participants perform two or more independent tasks either concurrently or alternating between the separate tasks, as in cases of task switching (see Koch et al., 2018, for a review of different forms of multitasking).

to activities critical for safe driving." By contrast, interdependent tasks support the same overall goal (e.g., maintaining vehicle control through speed and space management by looking at the speedometer and adjusting the accelerator to control vehicle speed and making steering inputs to stay in the lane of travel). Regan et al. (2011, p. 1975) defined the impairment from performing interdependent tasks (i.e., "driver misprioritized attention") as "insufficient or no attention to activities critical for safe driving brought about by the driver focusing attention on one aspect of driving to the exclusion of another, which is more critical for safe driving."

William James was one of the first to comment on the fact that the more similar two tasks are, the more they are likely to interfere with each other when they are performed concurrently (James, 1890). Early models of attention and dual-task performance (e.g., Kahneman, 1973; Navon & Gopher, 1979; Norman & Bobrow, 1975; Wickens, 1980, 1984, 2008) suggest that there is a limit on the attentional capacity of the human, that performance on a task is related to the attention allocated to that task, and that when people try to perform two or more attention-demanding tasks at the same time that there is a tradeoff between them such that one task prospers at the expense of the other. A conversational dyad produces a pattern of *resource reciprocity* such that the driving task competes for the same limited resources as the conversation task (Strayer, Biondi, Q6 et al., 2017; see also Castro et al., under review). By contrast, the *crosstalk hypothesis* suggests that dual-task interference occurs when two tasks use similar or overlapping information that come into conflict (e.g., Bergen et al., 2013; Navon & Miller, 1987; Pashler, 1994a). Crosstalk, sometimes described in the literature as "code conflict," has been shown to occur in perception, working memory, and in motor output. Importantly, Howard et al. (2020) used evidence accumulation modeling to show that multitasking differs from simply increasing task difficulty as the former decreases the rate of evidence accumulation *and* increases the response threshold whereas the latter decreases the rate of evidence accumulation but leaves the response threshold unchanged. This suggests that there is an added cost of multitasking that reflects a strategic adjustment to account for concurrently performing multiple tasks.

Salvucci and Taatgen (2008) developed a *threaded cognition* theory of multitasking to account for the costs incurred when people concurrently perform unrelated tasks such as driving while conversing on a smartphone. Based on the adaptive control of thought–rational (ACT-R) architecture, threaded cognition assumes that a serial cognitive processor coordinates the multiple task "threads" associated with currently active tasks. The theory instantiates an "exclusive-use" rule whereby requests for processing for other threads must wait until the completion of the current process. A procedural bottleneck arises when competition for the exclusive-use cognitive processor occurs. The exclusive-use rule causes behavior to be restricted to a single thread, a feature of cognition likely to minimize crosstalk between multiple task threads. Threaded cognition posits that task threads acquire processing resources in a "greedy" fashion and release them "politely." A greedy thread requests processing resources as soon as possible when they are needed, although the thread may have to wait its turn because of the exclusive-use rule. A polite thread releases resources for other threads as soon as its processing is no longer required. The task threads are prioritized so that the least recently processed thread receives priority. Because the cognitive processer operates on an exclusive-use rule, combining an attention-demanding secondary task with driving results in contention for cognitive processes and suboptimal driving performance (as well as suboptimal performance on the attention-demanding secondary task). Effectively, the current task thread locks out other threads from the central processor until processing on the current thread has completed.

The Detection Response Task

An early technique for measuring the attentional demands of a task is the secondary probe RT task (e.g., Posner & Boies, 1971; Q7 Posner & Keele, 1967). The procedure involved adding secondary task (e.g., a simple RT task with a light or tone as the stimulus and a simple button press as the response). Changes in the attentional demand of the primary task were inferred by the level of interference on a probe RT task. For example, an increase probe RT (or a decrease in the rate of detection of the stimulus—the Hit Rate) would be indicative of an increased in the attentional demands of the primary task. One advantage of the probe RT task is that it can provide a measure of the cognitive demands of the primary task even when measures of primary-task performance are difficult to obtain. Moreover, Posner and Keele (1967) found that the probe had little effect on the primary task, but that it was very sensitive to increased demand of the primary task.

The Detection Response Task (DRT; ISO 17488, 2016), is a standardized protocol for measuring cognitive demand from secondary-task engagement. Like the probe RT task developed by Posner and Keele (1967), the DRT protocol involves presenting a simple stimulus (e.g., a light or vibrotactile device) and requiring the participant to make a simple button press when they detect the stimulus. The DRT is useful because it is highly sensitive to changes in cognitive load, with increases in RT and decreases in Hit Rate as task demands increase. Because the DRT protocol involves presenting a simple stimulus every 3–5 s, a large and continuous amount of data are available for making inferences about changes in cognitive load. This attribute of the DRT is important because many other measures of performance are more difficult to obtain, occur infrequently or intermittently (e.g., brake RT), or are ambiguous in their interpretation. For example, lane-keeping performance can increase or decrease with increases in cognitive load depending on a variety of factors (e.g., Cooper et al., 2013; Medeiros-Ward et al., 2014).

The DRT is easy to administer and collect high-quality data. However, the biggest concern regarding the DRT is that it adds another task to the mix for participants to perform and that may alter the attentional demands and performance of the driver. Studies that have systematically compared driving performance with and without the DRT have found that it produces little or no interference with the other concurrent tasks, primarily because of the simple nature of the task. For example, Castro et al. (2019) found that the introduction of the DRT with an easily perceived visual stimulus and a simple response slightly degraded pursuit-tracking performance (the primary task). In a simulator study, Strayer et al. (2015) also found that subjective workload was not altered with the introduction of the DRT and Palada et al. (2019) found that the DRT did not significantly Q8 interfere with the primary task of classifying maritime ships as friend of foe. By contrast, Castro et al. (2019) reported that a visual stimulus that was difficult to perceive or a modification of the DRT task to a choice discrimination produced significantly greater interference on the primary task. The upshot is to keep the processing requirements of the DRT minimal and constant throughout the testing interval to facilitate the interpretation of the data.

Persistent Interference From Multitasking

In an on-road driving study, Strayer et al. (2016), see also Strayer, Q9 Biondi, et al. (2017) and Strayer, Cooper, et al. (2017) used the DRT to assess the residual costs after motorists had issued a voice command to tune the radio or to place a phone call. Multitasking costs persisted for much longer than what would be expected based on the PRP or task-switching literature (Koch et al., 2018). In fact, the residual costs followed a power function with the greatest costs observed immediately after the multitasking episode and persisted for up to 27 s after the secondary task had been completed.

Q10 Recently, Bowden et al. (2018) extended this work in a driving simulator by obtaining DRT measures when participants performed a concurrent driving task with a cognitive component, a cognitive and visual component, or a cognitive, visual, and manual component. In each case, DRT performance was significantly impaired following the secondary task with the initial DRT costs in the postsecondary task interval systematically increasing with the added complexity of the secondary task. Performance on the DRT did not return to single-task baseline levels for at least 60 s. Bowden et al. (2018), see also Jenness et al. (2015), also found postsecondary task impairments to the primary-task measures of lane keeping and driving speed; however, these residual costs had dissipated by 20 s after the secondary task had terminated.

There are at least two alternative interpretations of the residual costs reported by Bowden et al. (2018), Strayer et al. (2016), Q11 Strayer, Biondi, et al. (2017), and Strayer, Cooper, et al. (2017). The first, a *recovery of situational awareness* interpretation suggests that the residual costs are a consequence of the driver actively engaging in processes that restore the situational awareness that was lost while multitasking. When drivers multitask, their situational awareness can become compromised (e.g., Durso et al., 2007; Endsley, 1995, 2015; Horrey et al., 2006; Kass et al., 2007; Strayer & Fisher, 2016). The fidelity of a motorist's awareness of the driving situation governs their driving performance. In fact, Q12 Fisher and Strayer (2014), see also Endsley (1995), reported that a 5% decrease in a motorist's situational awareness doubles the relative risk of a crash. Under the situational awareness recovery interpretation, the residual costs observed in the postmultitasking interval reflect the beneficial processing related to restoring the degraded mental model of the motorist.

A second *proactive interference* interpretation, suggests that the residual costs reflect interference that gradually diminishes as the procedures and information used by the secondary task are gradually purged from working memory. When starting to multitask, motorists must load the secondary-task goals and procedures into working memory and then hold and manipulate that information (Baddeley & Logie, 1999; Engle, 2002). From the perspective of the threaded cognition (Salvucci & Taatgen, 2008), instead of the threads being "politely released" when no longer needed, the threads may be "sticky," gumming up the works until the information is purged from working memory (e.g., Poitras et al., 2020). Under the proactive interpretation, the residual costs observed in the postmultitasking interval reflect the costs of extraneous information in working memory that interferes with the primary task of driving.

Experiment 1

Experiment 1 sought to determine if the persistent costs occur in a controlled nondriving laboratory environment. Participants performed a simple pursuit-tracking task under two levels of secondary-task cognitive load. Performance was assessed both during and after the multitasking episode to determine if systematic persistent multitasking costs could be obtained in a simple nondriving context where there is little or no situational awareness to recover. Support for the recovery of situation awareness hypothesis would be obtained if the residual costs are abolished with the simple pursuit-tracking task. Support for the proactive interference hypothesis would be obtained if the residual costs persist even with the simple pursuit-tracking task.

Method

Participants

Thirty-two participants (16 men and 16 women) from the University of Utah and surrounding communities participated in the study. Participants ranged in age from 18 to 58 years ($\bar{x} = 25.8$). All reported normal neurological functioning, normal or corrected-to-normal visual acuity, held a valid driver's license and were fluent in English. G*Power (Faul et al., 2009) indicated that a paired samples *t*-test with 32 participants was sensitive to detect effects of $d = 0.55$ with 80% power (α—0.5, two-tailed).

Equipment

A pursuit-tracking task was displayed on a 106 cm diagonal Samsung monitor (1920 × 1,080 pixels) that was connected to a fixed-based driving simulator. The simulator was not used to provide a driving simulation; instead only the forward screen, steering wheel, and seat were utilized for the tracking task. Participants sat approximately 91 cm from the display. The tracking task entailed using the steering wheel to control a small triangle to align with a small yellow ball that moved horizontally across the screen. The yellow ball moved at a slow constant rate of 100 pixels/s and participant's triangle would remain green if near the target, and change to yellow and then red as it moved out of range of the target. A rotary encoder attached to the steering wheel recorded steering angle which was then fed into the host computer over a serial connection and translated into movement of the participants' triangle. Absolute deviations from the center of the participant's triangle to the center of the target ball were recorded in pixels and stored for analysis.

The DRT protocol involved attaching a vibrotactile device to the participant's left collarbone and having participants respond by pressing a microswitch that was attached to the index of the left hand against the steering wheel. Following the ISO guidelines (2016), the vibrotactile device emitted a small vibration stimulus, similar to a vibrating cell phone, every 3–5 s (i.e., a rectangular distribution of interstimulus intervals between 3 and 5 s) and stimulation lasted for 1 s or until the participant pressed the microswitch.

The DRT software recorded participants' reaction time, with millisecond accuracy. Following ISO standards, a response was considered a hit if the RT was between 100 and 2,500 ms and a miss after 2,500 ms (ISO 17488, 2016). When multitasking, participants counted backwards for 20 s. Because there was often a slight pause between the counting prompt and participants becoming fully engaged with the counting task, the average of the last 5–6 s of on-task performance was used to calculate on-task RT and Hit Rate.

Procedure

Upon arrival, participants completed a consent document and basic demographic questionnaire. Participants were familiarized with the tracking task, DRT, and counting backwards by 1's and 3's. For the counting backwards tasks, participants counted backwards by either 1's or 3's from a randomly selected number between 200 and 999 (Pellecchia et al., 2005). Counting backwards by 1's or 3's predictably alternated throughout the experiment (see Figure 1). Participants' responses were recorded to calculate counting rate and accuracy.

Participants began the study with a 3-min baseline in which they performed the tracking task and responded to the DRT stimuli. After the 3-min baseline, the researcher provided a prompt number to participants (e.g., "By 1's, 213"), and participants would count backwards from the indicated number for 20 s while continuing the tracking task and responding to the DRT. After 20 s, the researcher would say, "Stop," and participants would continue the tracking task and DRT for a 30 s recovery period before the researcher would provide the next prompt. After the first block, participants would take a short break followed by an identical block of baseline followed by 20 more numbers from which to count backwards. Over the course of the study, each participant experienced 90 DRT trials in the baseline periods, 400 DRT trials in the on-task period, and 600 DRT trials in the off-task period.

Results

Counting Backwards

Auditory recordings of participants' counting backwards performance were scored both for total number completed within the 20 s (rate) and for accuracy (n = 29, as three audio files were corrupted). Counting accuracy was determined on an item-by-item basis, so an error on one "trial" would not count against subsequent counts in the sequence. Counting backwards by 1's, participants counted an average of .90 numbers per second (i.e., 17.9 numbers/20 s). While counting backwards by 3's, participants' speed significantly slowed to a rate of .51 numbers per second (i.e., 10.3 numbers/20 s), $t(28) = 12.34$, $p < .001$, $d = 1.60$. Additionally, participants' accuracy significantly differed between counting backwards by 1's (98.1%) and 3's (94.3%), $t(28) = 2.58$, $p = .015$, $d = 0.32$). Thus, participants decreased both their counting speed and they were less accurate when counting backwards by 3's compared to 1's.

Pursuit-Tracking Task

The RMSE steering error was 2.94 (SD = 1.54), which is indicative of a high degree of tracking accuracy (Castro et al., 2019). Q13 Participant's performance between the first (M = 2.96, SD = 2.36) and second (M = 2.92, SD = 1.55) blocks did not significantly differ, $t(30) = 0.76$, $p = .94$.

DRT

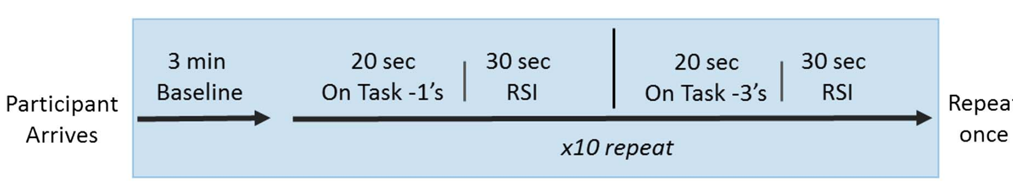
On-task RT and Hit Rate were calculated based on the average of the last two bins (i.e., 18 and 21 s in the on-task interval). Participants' average RT was 332.3 ms in the baseline drive, 572.1 ms while on-task counting backwards by 1's, 686.8 ms while on-task counting backwards by 3's. Participants' average Hit Rate was 99.4% in the baseline drive, 93.7% while on-task counting backwards by 1's, 87.1% while on-task counting backwards by 3's. RT was significantly faster when counting backwards by 1's than 3's, $t(31) = 4.43$, $p < .001$, $d = 0.54$, and Hit Rate was significantly higher when counting backwards by 1's than 3's, $t(31) = 4.19$, $p < .001$, $d = 0.53$.

Based on a visual inspection of the data, a series of contrasts were conducted on bins 3–18 in the off-task interval to determine the residual costs relative to baseline. There were approximately 30 observations for each participant at each of the bins in the analysis. The DRT RT data, grouped into 3-s bins relative to the onset and Q14 offset of the counting task off-task, are plotted in Figure 2 and the results of the planned comparisons are presented in Table 1. Planned comparisons between the single-task baseline and the consecutive 3-s off-task bins were significant for all comparisons between bins 3 and 18 for both backwards counting tasks. A similar comparison between the average of the last three bins in the on-task interval (i.e., bins 12–18) and the consecutive 3-s off-task bins found no significant difference for bins 3 and 6, but this difference was significant for bins 9–18 for both backwards counting tasks.

The DRT Hit Rate data, grouped into 3-s bins relative to the onset and offset of the counting task, are plotted in Figure 3 and the results of the planned comparisons are presented in Table 2. Planned comparisons between the single-task baseline and the consecutive 3-s off-task bins were significant for bins 3 and 6, but not for bins 9–18 for both backwards counting tasks. A similar comparison between the average of the last three bins in the on-task interval (i.e., bins 12–18) and the consecutive 3-s off-task bins found no significant difference for bins 3, 6, and 12 but this difference was significant for bins 9, 15, and 18 in the backwards counting by 1's task. For the backwards counting by 3's task, no significant

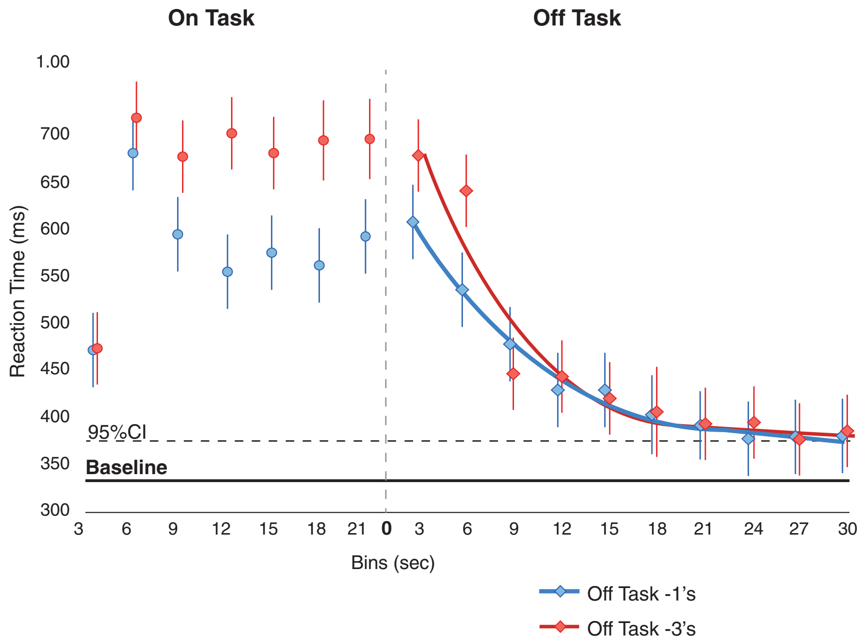
Figure 1

Protocol Timing for One of the Two Identical Blocks



Note. See the online article for the color version of this figure.

Figure 2
Q15 *Participants' DRT RT in Experiment 1*



COLOR

Note. The gray vertical dashed line indicates the point at which participants ceased counting. The solid blue and red lines represent the best-fitting power functions of off-task RT for 1's (R^2 = .99) and 3's (R^2 = .93), respectively.

difference was found for bins 3 and 6, but the difference was significant for bins 9–18.

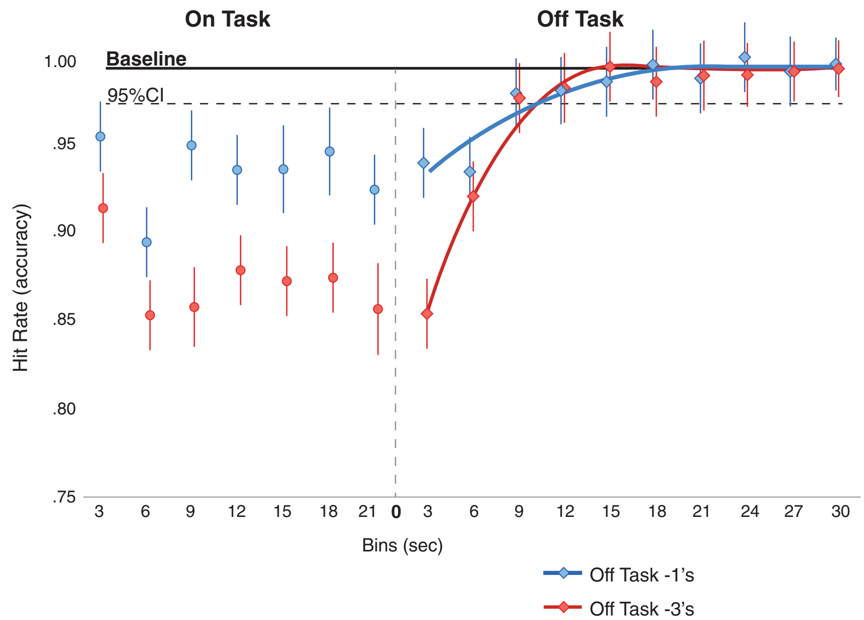
Discussion

Dual-task performance involved a simple pursuit-tracking task that was paired with two variants of a backwards counting task. Both RT and Hit Rates to the DRT were sensitive to on-task and off-task secondary-task difficulty. Importantly, RT differences were observed throughout the off-task interval, with greater costs observed immediately after counting stopped. These residual costs were greater for the counting backwards by 3's than 1's. Taken together, these data establish that the residual costs observed in the complex on-road driving environment reported by Strayer et al. (2016), Strayer, Biondi, et al. (2017), and Strayer, Cooper, et al. (2017) can be obtained in a simple pursuit-tracking task, that these Q21 effects are obtained for both RT and Hit Rates, and that they are sensitive to secondary-task load. Due to the simple nature of the

Table 1
Experiment 1: Contrasts Between Baseline or On-Task and Off-Task DRT RT Q17

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(31)	*p*	*d*	*t*(31)	*p*	*d*
	332.26	54.03	572.08	135.29	686.77	162.943						
			Off-task −1's									
Bin 3			607.34	112.66			14.91	.001	1.83	2.43	ns	0.30
Bin 6			542.77	85.70			11.41	.001	1.40	−1.03	ns	−0.13
Bin 9			470.48	87.80			7.49	.001	0.92	−4.91	.001	−0.60
Bin 12			421.33	77.41			4.83	.001	0.59	−7.55	.001	−0.92
Bin 15			424.77	95.42			5.01	.001	0.61	−7.36	.001	−0.90
Bin 18			401.20	85.15			3.74	.001	0.46	−8.63	.001	−1.06
					Off-task −3's							
Bin 3					677.82	134.95	14.382	.001	1.90	−0.29	ns	−0.04
Bin 6					640.28	134.41	12.802	.001	1.69	1.70	ns	−0.23
Bin 9					445.53	84.22	4.715	.001	0.62	−8.91	.001	−1.20
Bin 12					443.49	84.18	4.630	.001	0.61	−8.98	.001	−1.21
Bin 15					416.83	70.91	3.520	.003	0.47	−9.97	.001	−1.34
Bin 18					399.90	87.22	2.816	.032	0.37	−10.60	.001	−1.43

Figure 3
Participants' DRT Hit Rate in Experiment 1



COLOR

Note. The gray vertical dashed line indicates the point at which participants ceased counting. The solid blue and red lines represent the best-fitting power functions of off-task Hit Rate for 1's (R^2 = .88) and 3's (R^2 = .98), respectively. DTH = detection response task.

Q18

tracking task, these data are more consistent with the proactive interference interpretation than the recovery of situational awareness interpretation.[2]

Experiment 2

The objective of Experiment 2 was to evaluate the residual costs in a controlled driving simulator where both the driving difficulty and secondary-task load were manipulated. It was predicted that the residual costs would be sensitive to both factors. Additionally, we collected eye-tracking measures to determine if there were changes in pupil diameter and the visual scanning patterns following a multitasking episode. The pupil dilation response (PDR) reflects increases in arousal and effort in response to increases in cognitive demand (e.g., Ahern & Beatty, 1979; Beatty & Lucero-Wagoner, 2000; Kahneman & Beatty, 1966; Silcox & Payne, 2021; Zekveld & Kramer, 2014). The PDR has been linked to phasic activity in the locus coeruleus norepinephrine system (Joshi & Gold, 2020; LoTemplio et al., 2021). If the residual costs in the DRT observed in Experiment 1 reflect lingering levels of cognitive load, then changes in the PDR across the off-task interval should approximate the pattern observed with the DRT.

We also looked for evidence in the eye-tracking measures of enhanced visual scanning in the off-task interval that would suggest an oversampling of the primary task (relative to single-task baseline levels). According to the recovery of situation awareness hypothesis, participants may try to update their mental model of the driving environment after a multitasking event. A pattern of increased visual scanning (e.g., scanning the periphery to update the status of adjacent vehicles on the roadway or looking at the instrument cluster to update the status of their driving speed) to areas that were neglected while multitasking would suggest an active process to recover situation awareness.

Method

Participants

Forty-seven participants (22 men and 25 women) from the University of Utah and surrounding communities participated in the study. Participants ranged in age from 18 to 52 years ($\bar{x}$ = 23.0). All reported normal neurological functioning, normal or corrected-to-normal visual acuity and hearing, held a valid driver's license, and were fluent in English. Participants' years of driving experience ranged from 1 to 27 years ($\bar{x}$ = 6.8). Participants reported driving an average of between 0 and 300 miles per week ($\bar{x}$ = 107.6). All participants owned a cell phone and 55.3% reported that they used their phones regularly while driving. G*Power indicated that 45 participants were needed to discern differences with 80% power with an effect size f of 0.33, an α level of .05 (Faul et al., 2009).

[2] An alternative interpretation of these data is that the monotonic RT costs in the off-task interval are an artifact of averaging over step functions from different participants. While such an interpretation does predict a smooth decrease in mean RT, it also predicts an inverted *U*-shaped pattern in the RT standard deviation caused by the mixture of single-task and dual-task performance. Inspection of Table 1; however, shows a gradual decrease in the RT standard deviation, suggesting that the data are not an artifact of averaging over step functions from different participants.

Table 2
Experiment 1: Contrasts Between Baseline or On-Task and Off-Task DRT Hit Rates Q19

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(31)	*p*	*d*	*t*(31)	*p*	*d*
	99.40	1.63	93.68	9.83	87.12	17.65						
			Off-task −1's									
Bin 3			93.70	22.46			3.69	.001	0.79	−0.22	ns	−0.04
Bin 6			93.35	24.15			4.09	.001	0.87	−.59	ns	−0.12
Bin 9			98.54	12.65			0.58	ns	0.12	2.69	.047	0.53
Bin 12			98.50	12.91			0.64	ns	0.14	2.63	ns	0.52
Bin 15			98.89	12.19			0.25	ns	0.05	2.99	.019	0.59
Bin 18			99.74	7.10			0.28	ns	0.06	3.49	.003	0.69
					Off-task −3's							
Bin 3					85.87	16.41	6.85	.001	1.34	−0.07	ns	−0.12
Bin 6					91.45	7.35	3.62	.002	0.71	2.18	ns	0.40
Bin 9					97.59	5.20	0.88	ns	0.17	4.62	.001	0.84
Bin 12					98.01	6.07	0.60	ns	0.12	4.87	.001	0.88
Bin 15					99.74	4.97	0.18	ns	0.04	5.56	.001	1.01
Bin 18					98.64	4.84	0.32	ns	0.06	5.12	.001	0.93

Q20 *Note.* DRT = Detection Response Task.

Equipment

A DS-600 fixed-base driving simulator was used for this study. Two driving scenarios were designed to provide both a low traffic (Easy) and a dense traffic (Hard) driving environment. The driving scenarios consisted of a 19-mile loop that alternated between two and three lanes following the suggestions of Carsten et al. (2005) and Normark et al. (2009) regarding road types, speed limits, traffic type, and other driving relevant objects. Driving speeds were regulated between 55 and 65 mph by speed limit signs in both scenarios. Programmed trigger points in both scenarios controlled surrounding traffic to match with the drivers' progression through the loop, creating the *irregular-flow* traffic (Drews et al., 2008). Q22

Traffic in the Easy scenario consisted of approximately 500 vehicles per lane per hour (traffic density = 19 vehicles/mile), which gave the impression of light traffic, mostly in the opposite direction of the highway. Traffic in the Hard scenario was operationalized as a traffic flow of approximately 1,700 vehicles per lane per hour (traffic density = 28 vehicles/mile). The greater number of vehicles in this scenario often slowed traffic in the right lane at three- to two-lane transition points. The traffic back up was noticeable and easily avoided if anticipated.

Q23 Eye movements were recorded using a Seeing Machines Fovio Eye Tracker. The Fovio was a dash-mounted device that did not interfere with participants' view and allowed free range of head and eye movements. The eye tracker was calibrated at the beginning of each of the drives and participants sat an average of 60 cm from the Fovio tracker. The DRT equipment was identical to Experiment 1.

Procedure

Upon arrival for their first appointment, participants completed a brief demographic questionnaire and warm-up driving scenario. Participants completed a total of four driving scenarios between two appointments. The order of conditions was counterbalanced between beginning the study in either the Easy or Hard driving scenarios, and each appointment consisted of one Easy and one Hard drive.

Participants were trained on responding to the DRT stimuli and on counting backwards by 1's and 3's. Participants were instructed to remain in the right-most lane of the highway except when necessary to pass traffic that was traveling under the posted speed limit. Participants began each of the drives with a 3-min baseline period wherein they were just driving and responding to the DRT. After the 3 min, the researcher began providing prompt numbers to the participants in the same manner as Experiment 1 (e.g., "By 3's, 272"). Participants counted backwards for 20 s before the researcher said, "Stop," and then continued driving and responding to the DRT for a 45 s recovery period before the next prompt. After 10 repetitions of both counting backwards by 1's and by 3's (approximately 26 min), participants took a short break, and then resumed the second drive (see Figure 4). Throughout the total duration of the study, each participant experienced an average of 180 DRT trials during the baseline periods, 400 on-task, and 900 during the recovery period. There were 60 observations for each participant at each of the bins in the analysis of the off-task data. At their second appointment, participants completed just the remaining two drives.

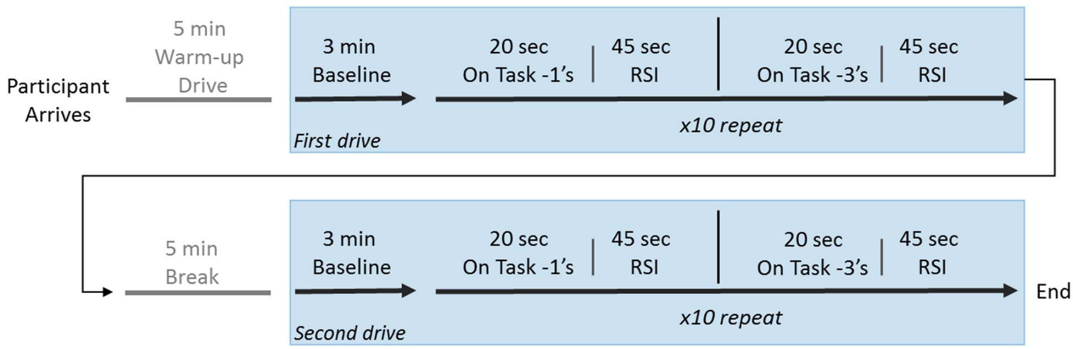
Results

Counting Backwards

Recordings of participants' counting backwards performance were scored both for a total number of steps backwards completed within the 20 s (rate) and for accuracy ($n = 46$, as one audio file was corrupted). Table 3 documents both the rate and accuracy for Easy and Hard drives and counting backwards by 1's and 3's.

A two-factor repeated-measures ANOVA on the accuracy revealed a significant effect of counting task difficulty, $F(1, 45) = 25.08$, $p <$ .001, $\eta_p^2 = .36$, such that accuracy when counting backwards by 3's was significantly lower than when counting by 1's. Neither the effect of driving environment ($p = .57$), nor the counting difficulty by driving environment interaction ($p = .89$) was significant.

Figure 4
Protocol for One Driving Appointment Consisting of Two 26-Min Drives With Easy and Hard Driving Environments Counterbalanced



ONLINE COLOR

Note. See the online article for the color version of this figure.

Driving Performance

In the Easy driving environment, the average lane position was −.09 ($SD = .41$, indicative of remaining in the center of the lane) and the average speed was 61.9 m/hr ($SD = 5.49$). Demonstrating the greater driving demand of the Hard driving environment, average lane position was −.13 ($SD = .41$, indicating participants drove slightly left of the center of their lane) and their speed decreased to 56.4 m/hr ($SD = 9.05$). A paired samples *t*-test demonstrated significant differences between Easy and Hard driving environments for both lane position, $t(46) = 2.65$, $p = .01$, $d = 0.42$ and speed, $t(46) = 8.89$, $p < .001$, $d = 1.09$.

DRT

Participants' average RT for each period of the drive is presented in Figure 5. A two-factor repeated-measures ANOVA on the RT revealed a significant effect of counting difficulty, $F(1, 45) = 110.14$, $p < .001$, $\eta_p^2 = .71$, and of driving environment, $F(1, 45) = 4.53$, $p = .04$, $\eta_p^2 = .09$; however, the counting difficulty by driving environment interaction was not significant ($p = .12$). The results of the planned comparisons are presented in Tables 4 and 5. Planned comparisons between the single-task baseline and the consecutive 3-s off-task bins were significant for all comparisons between bins 3 and 18 for both backwards counting tasks and for both Easy and Hard driving conditions. A similar comparison between the average of the last three bins in the on-task interval (i.e., bins 12–18) and the consecutive 3-s off-task bins found no significant difference for bins 3–9, but this difference was significant for bins 12–18 for both the Easy and Hard driving conditions when participants were counting backwards by 1's. The same pattern was obtained when participants were counting backwards by 3's, with the exception that bin 9 was also significantly different from the on-task average for both Easy and Difficult driving conditions.

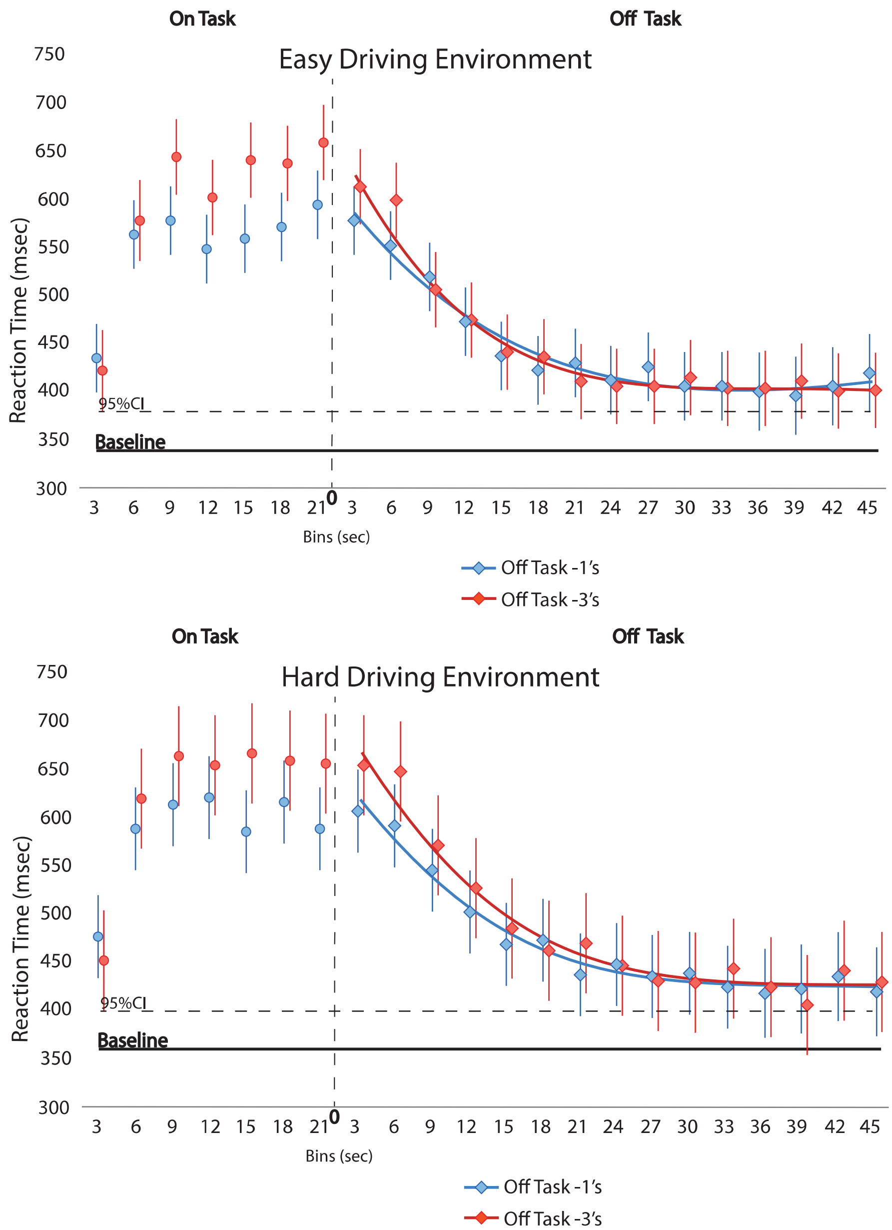
Participants' average Hit Rate on the DRT for each period of the drive is presented in Figure 6. The baseline periods between the Easy and Hard driving scenarios were significantly different, $t(46) = -2.44$, $p = .019$, $d = 0.39$, indicating that participants experienced greater difficulty in the Hard driving condition. The results of the planned comparisons are presented in Tables 6 and 7. Planned comparisons between the Easy driving single-task baseline and the consecutive 3-s off-task bins were significant for all comparisons for bins 3–12, but the difference was not significant for bins 15 and 18 for either backwards counting tasks. Planned comparisons between the Hard driving single-task baseline and the consecutive 3-s off-task bins were significant for all comparisons for bins 3–15, but the difference was not significant for bin 18 when counting backwards by 1's. When participants were counting backwards by 3's, the pattern was the same with the exception that the difference at bin 15 was not significant.

A comparison between the average of the last three bins in the on-task interval (i.e., bins 12–18) and the consecutive 3-s off-task bins in the Easy driving condition found no significant difference for bins 3–15, but this difference was significant for bin 18 when participants were counting backwards by 1's. When participants were counting backwards by 3's, the pattern was the same with the exception that the difference at bins 15 and 18 was significant. In the Hard driving condition, there were no significant difference for bins 3–9, but the difference was significant for bins 12–18 when participants were counting backwards by 1's. When participants were counting backwards by 3's, the pattern was the same with the exception that the difference at bins 15 and 18 was significant.

Table 3
Q24 *Counting Backwards Performance*

	Easy drive				Hard drive			
	Rate (#/s)	*SD*	Accuracy (%)	*SD*	Rate (#/s)	*SD*	Accuracy (%)	*SD*
Q25 −1's	.79	*.27*	98.9	*.01*	.76	*.24*	99.1	*.01*
−3's	.53	*.17*	96.0	*.05*	.51	*.16*	96.2	*.04*

Figure 5
Participants' DRT RT in Experiment 2



Note. The gray vertical dashed line indicates the point at which participants ceased counting. In the Easy driving environment (top panel), the solid blue and red lines represent the best-fitting power functions of off-task RT for 1's (R^2 = .97) and 3's (R^2 = .97), respectively. In the Hard driving environment (bottom panel), the solid blue and red lines represent the best-fitting power functions of off-task RT for 1's (R^2 = .97) and 3's (R^2 = .97), respectively.

Pupillometry

An eye tracker was utilized to measure the PDR of participants across conditions. PDR has been found to increase with increases in cognitive load (e.g., Ahern & Beatty, 1979; Beatty & Lucero-Wagoner, 2000; Kahneman & Beatty, 1966; Silcox & Payne, 2021; Zekveld & Kramer, 2014). Three participant's eye-tracking data were removed due to a recording malfunction, leaving 43 participants. Data processing was conducted with the Saccades R package (R Core Team, 2019; von der Malsburg, 2015). For the conducted Trackloss analysis, 15% (*SD* = 8%) of data per participant was removed on average. Blinks were also identified through the Trackloss procedure and removed.

We assessed how the difficulty of the counting task (i.e., 1's vs. 3's) and the driving environment (i.e., Easy vs. Hard) affected pupil

Table 4
Contrasts Between Baseline and Off-Task DRT RT for the Easy Drive Q26

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	331.57	84.57	551.30	79.08	622.87	111.66						
			Off-task −1's									
Bin 3			578.11	106.51			13.22	.001	1.69	1.64	ns	0.22
Bin 6			548.40	124.25			11.63	.001	1.49	0.06	ns	0.01
Bin 9			505.30	79.42			9.31	.001	1.19	−2.20	ns	−0.30
Bin 12			464.03	59.32			7.10	.001	0.91	−4.37	.001	−0.59
Bin 15			430.79	65.23			5.32	.001	0.68	−6.11	.001	−0.82
Bin 18			419.56	69.15			4.72	.001	0.61	−6.70	.001	−0.90
					Off-task −3's							
Bin 3					594.67	117.40	13.81	.001	1.55	−1.57	ns	−0.19
Bin 6					598.39	116.57	14.01	.001	1.36	−1.38	ns	−0.17
Bin 9					498.41	72.87	8.76	.001	1.09	−6.39	.001	−0.76
Bin 12					467.02	67.38	7.11	.001	0.83	−7.96	.001	−0.95
Bin 15					436.94	67.45	5.53	.001	0.62	−9.46	.001	−1.13
Bin 18					430.71	67.40	5.20	.001	0.55	−9.77	.001	−1.17

diameter. Participants' average pupil diameter for each period of the drive is presented in Figure 7. A two-factor repeated-measures ANOVA of pupil diameter demonstrated a significant effect of counting task difficulty, $F(1, 42) = 91.42$, $p < .001$, $\eta_p^2 = .42$, and driving environment $F(1, 42) = 67.80$, $p < .001$, $\eta_p^2 = .35$. There was no significant interaction between the driving environment and counting task difficulty $F(1, 42) = .32$, $p = .57$, $\eta_p^2 < .01$. Planned comparisons between the Easy driving single-task baseline and the consecutive 3-s off-task bins were significant for all comparisons for bins 3–15 for counting backwards by 1's and for bins 3–9 when counting backwards by 3's (see Table 8). Planned comparisons between the Hard driving single-task baseline and the consecutive 3-s off-task bins were significant for bins 3–15 when counting backwards by 1's and significant for all bins when counting backwards by 3's (see Table 9). A comparison between the average of the last three bins in the on-task interval (i.e., bins 12–18) and the consecutive 3-s off-task bins found significant difference at all bins for both Easy and Hard driving conditions and for both counting tasks.

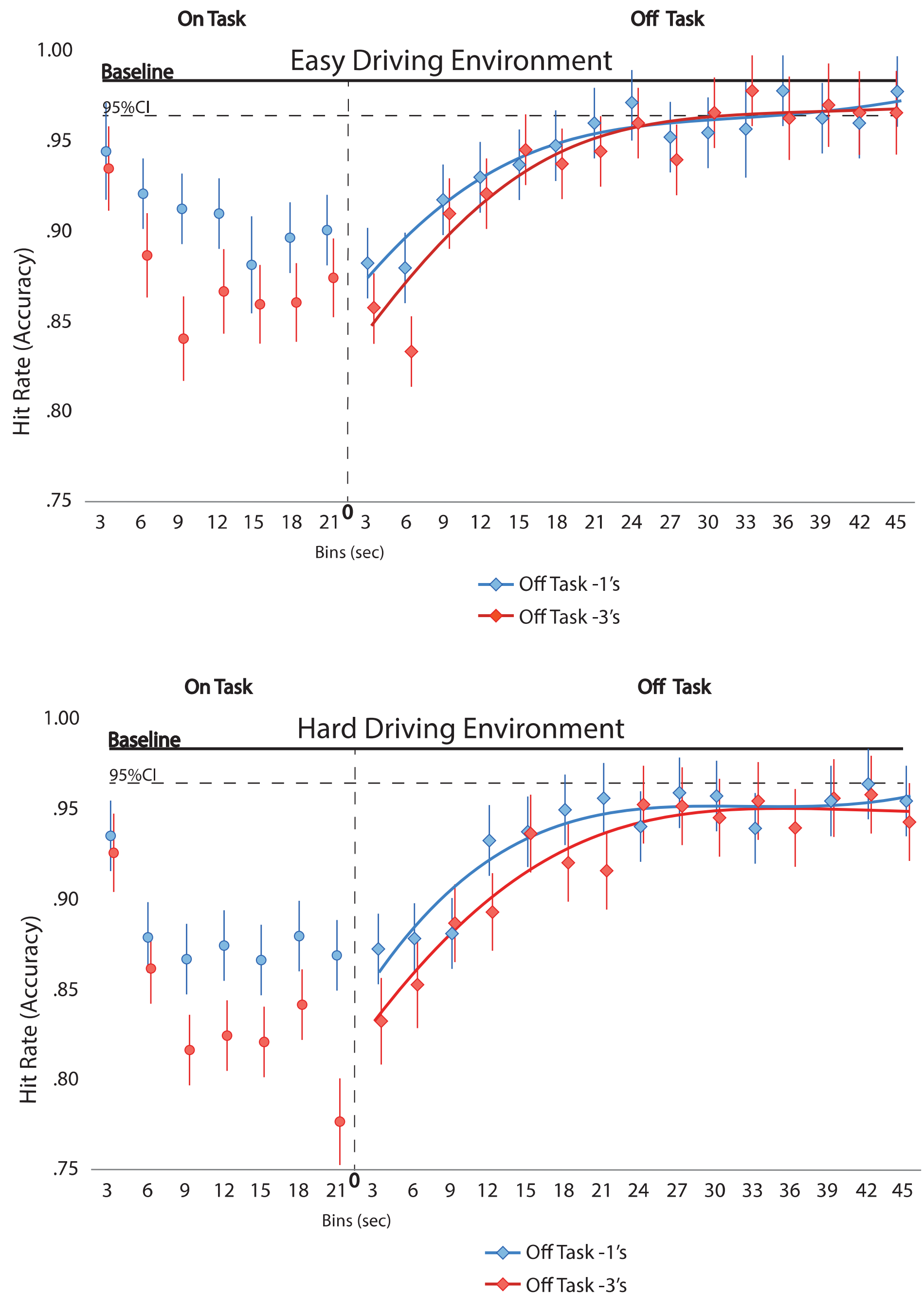
Eye Tracking

The eye tracker was also used to determine if there were changes in visual scanning in the off-task interval that might be indicative of oversampling of the driving environment to recover situation awareness that was lost in the on-task interval. We extracted two measures: The percentage of time the participant's gaze was fixated on the forward roadway regions of interest ROI (the area outlined in red in Figure 8, encompassing 10.48% of the view from the scene camera) and the percentage of time participant's gaze was fixated on the instrument panel (the area outlined in yellow in Figure 8, encompassing 5.86% view from the scene camera).

Table 5
Contrasts Between Baseline and Off-Task DRT RT for the Hard Drive Q27

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	354.11	72.3	598.17	168.74	654.18	213.84						
			Off-task −1's									
Bin 3			596.57	147.90			10.75	.001	1.51	0.35	ns	.05
Bin 6			586.59	110.64			10.31	.001	1.33	−0.08	ns	−0.011
Bin 9			536.92	115.92			8.11	.001	1.07	−2.25	ns	−0.320
Bin 12			494.58	70.84			6.23	.001	0.81	−4.09	.001	−0.583
Bin 15			460.13	89.45			4.70	.001	0.61	−5.59	.001	−0.797
Bin 18			472.32	88.64			5.24	.001	0.54	−5.06	.001	−0.721
					Off-task −3's							
Bin 3					631.61	162.25	10.71	.001	1.31	−0.35	ns	−0.05
Bin 6					638.06	154.54	10.96	.001	1.15	−0.11	ns	−0.02
Bin 9					566.42	118.16	8.20	.001	0.92	−2.83	.03	−0.39
Bin 12					518.58	93.03	6.35	.001	0.70	−4.65	.001	−0.64
Bin 15					479.76	97.74	4.85	.001	0.53	−6.13	.001	−0.85
Bin 18					460.63	83.00	4.11	.001	0.47	−6.85	.001	−0.95

Figure 6
Participants' DRT Hit Rate in Experiment 2



Note. In the Easy driving environment (top panel), the solid blue and red lines represent the best-fitting power functions of off-task Hit Rate for 1's ($R^2 = .92$) and 3's ($R^2 = .88$), respectively. In the Hard driving environment (bottom panel), the solid blue and red lines represent the best-fitting power functions of off-task Hit Rate for 1's ($R^2 = .86$) and 3's ($R^2 = .93$), respectively. DTH = detection response task.

We first assessed how the difficulty of the counting task (i.e., 1's vs. 3's) and the driving environment (i.e., Easy vs. Hard) affected the amount of time looking to the forward roadway. A two-factor repeated-measures ANOVA demonstrated a significant effect of counting task difficulty, $F(1, 42) = 7.91$, $p = .007$, $\eta_p^2 = .06$, but not driving environment $F(1, 42) = .836$, $p = .37$, $\eta_p^2 = .04$. The interaction between counting difficulty and driving environment was not significant, $F(1, 42) = .06$, $p = .81$, $\eta_p^2 < .01$. This pattern

Table 6
Contrasts Between Baseline and Off-Task DRT Hit Rate for the Easy Drive Q28

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	98.45	7.26	89.81	8.53	87.47	13.71						
			Off-task −1's									
Bin 3			89.05	9.98			5.25	.001	0.83	−0.42	ns	−0.06
Bin 6			88.16	10.92			5.74	.001	0.91	−0.92	ns	−0.13
Bin 9			92.80	7.27			3.16	.010	0.50	1.72	ns	0.25
Bin 12			93.25	6.24			2.90	.024	0.46	1.98	ns	0.29
Bin 15			93.86	4.44			2.56	ns	0.40	2.33	ns	0.34
Bin 18			94.93	8.54			1.96	ns	0.31	2.93	.02	0.43
					Off-task −3's							
Bin 3					87.13	9.57	5.70	.001	0.89	−0.15	ns	−0.02
Bin 6					83.78	11.19	7.39	.001	1.16	−1.81	ns	−0.27
Bin 9					91.52	9.55	3.49	.003	0.55	2.03	ns	0.30
Bin 12					92.46	10.65	3.02	.017	0.47	2.50	ns	0.37
Bin 15					94.79	6.80	1.84	ns	0.29	3.66	.002	0.54
Bin 18					93.89	9.29	2.30	ns	0.36	3.21	.009	0.47

Note. DRT = Detection Response Task.

shows that participants concentrated their gaze on the forward roadway in the on-task intervals (e.g., Reimer, 2009; Wang et al., 2014). Planned comparisons between the Easy driving single-task baseline and the consecutive 3-s off-task bins were not significant for any of the bins when counting backwards by 1's and was significant for only bin 9 when counting backwards by 3's (see Table 10). Planned comparisons between the Hard driving single-task baseline and the consecutive 3-s off-task bins were not significant for any of the bins when counting backwards by 1's and was significant for bins 15 and 18 when counting backwards by 3's (see Table 11). A comparison between the average of the last three bins in the on-task interval (i.e., bins 12–18) and the consecutive 3-s off-task bins in the Easy driving condition found significant difference at bins 3 and 18 when participants were counting backwards by 1's and a significant difference at bin 3 when participants were counting backwards by 3's. A comparison between the average of the last three bins in the on-task interval (i.e., bins 12–18) and the consecutive 3-s off-task bins in the Hard driving condition found significant difference for all conditions when participants were counting backwards by 1's or 3's (Figure 9). Q35

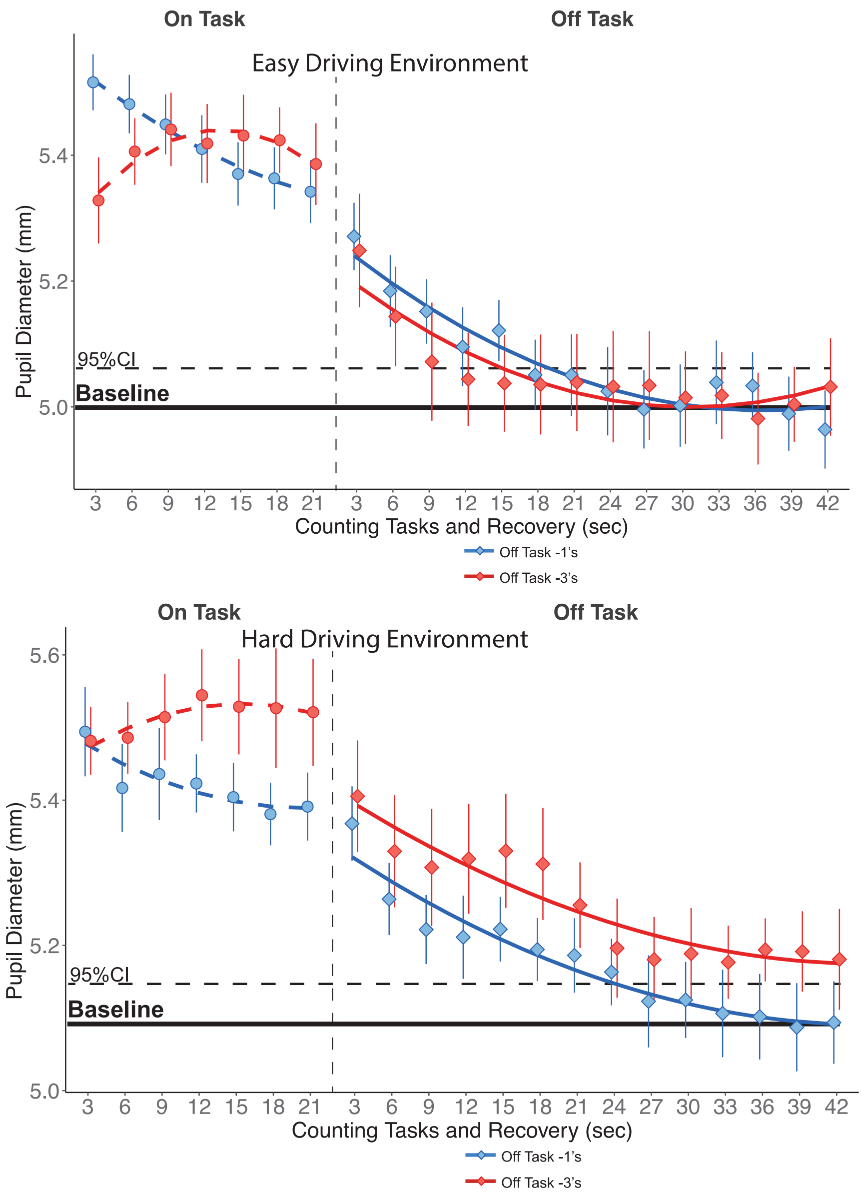
Next, we assessed how the difficulty of the counting task (i.e., 1's vs. 3's) and the driving environment (i.e., Easy vs. Hard) affected glances to the instrument cluster. A two-factor repeated-measures ANOVA demonstrated a significant effect of counting task difficulty, $F(1,\ 42) = 8.14$, $p = .007$, $\eta_p^2 = .28$, but not driving environment $F(1,\ 42) = 1.27$, $p = .27$, $\eta_p^2 = .14$. There was no significant interaction between counting difficulty and driving

Table 7
Contrasts Between Baseline and Off-Task DRT Hit Rate for the Hard Drive Q29

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	98.44	2.2	87.7	15.1	81.0	20.5						
			Off-task −1's									
Bin 3			87.79	6.25			5.889	.001	0.97	0.27	ns	0.04
Bin 6			88.14	9.82			5.696	.001	0.94	0.47	ns	0.07
Bin 9			88.62	9.95			5.431	.001	0.90	0.70	ns	0.11
Bin 12			93.18	8.44			2.906	.024	0.48	3.07	.014	0.48
Bin 15			93.51	7.23			2.723	.042	0.45	3.24	.008	0.51
Bin 18			95.10	6.92			1.843	ns	0.30	4.07	.001	0.64
					Off-task −3's							
Bin 3					84.90	8.84	6.72	.001	0.81	0.14	ns	0.02
Bin 6					85.68	13.24	6.33	.001	0.79	0.52	ns	0.07
Bin 9					89.25	9.95	4.56	.001	0.75	2.26	ns	0.32
Bin 12					89.63	7.59	4.37	.001	0.40	2.44	ns	0.35
Bin 15					93.89	6.80	2.26	ns	0.38	4.51	.001	0.64
Bin 18					92.42	6.76	2.99	.019	0.25	3.80	.001	0.54

Note. DRT = Detection Response Task.

Figure 7
Pupil Diameter Response in Experiment 2



COLOR

Note. The gray vertical dashed lines indicate the point at which participants ceased counting. In the Easy driving environment (top panel), the solid blue and red lines represent the best-fitting power functions of off-task pupil diameter for 1's (R^2 = .78) and 3's (R^2 = .84), respectively. In the Hard driving environment (bottom panel), the solid blue and red lines represent the best-fitting power functions of off-task pupil diameter for 1's (R^2 = .86) and 3's (R^2 = .82), respectively.

environment, $F(1,42) = 1.99$, $p = .17$, $\eta_p^2 < .01$. Planned comparisons between the Easy driving single-task baseline and the consecutive 3-s off-task bins were not significant for any of the bins when counting backwards by 1's or 3's (see Table 12). Planned comparisons between the Hard driving single-task baseline and the consecutive 3-s off-task bins were not significant for any of the bins when counting backwards by 1's and was significant for only bin 12 when counting backwards by 3's (see Table 13). A comparison between the average of the last three bins in the on-task interval (i.e., bins 12–18) and the consecutive 3-s off-task bins in the Easy driving condition found significant difference at all bins for Easy and Hard driving conditions when participants were counting backwards by 1's. A comparison between the average of the last three bins in the on-task interval (i.e., bins 12–18) and the consecutive 3-s off-task bins when participants were counting backwards by 3's found significant differences at bins 3, 6, and 12 in the Easy driving condition and at bins 3 and 9 in the Hard driving condition (Figure 10). Q38

Table 8
Contrasts Between Baseline and Off-Task Pupil Diameter for the Easy Drive Q30

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	5.00	0.15	5.38	0.09	5.42	0.10						
			Off-task −1's									
Bin 3			5.18	0.10			7.15	.001	0.27	−9.19	.001	−0.31
Bin 6			5.12	0.11			4.45	.001	0.17	−12.14	.001	−0.41
Bin 9			5.10	0.08			3.79	.001	0.14	−12.90	.001	−0.44
Bin 12			5.09	0.09			3.41	.005	0.13	−13.31	.001	−0.45
Bin 15			5.10	0.09			3.93	.001	0.15	−12.80	.001	−0.44
Bin 18			5.05	0.10			1.68	ns	0.04	−15.29	.001	−0.52
					Off-task −3's							
Bin 3					5.19	0.12	6.60	.001	0.27	−4.89	.001	−0.22
Bin 6					5.14	0.09	4.59	.001	0.19	−6.72	.001	−0.30
Bin 9					5.09	0.08	2.81	.031	0.11	−8.29	.001	−0.37
Bin 12					5.08	0.10	2.36	ns	0.10	−8.73	.001	−0.38
Bin 15					5.08	0.10	2.26	ns	0.09	−8.47	.001	−0.37
Bin 18					5.08	0.12	2.28	ns	0.09	−8.45	.001	−0.37

General Discussion

Multitasking, the act of performing two or more functionally independent tasks with separate goals, stimuli, and responses, is a common occurrence in everyday life. The objective of this research was to examine a hidden cost of multitasking. We used the DRT, an ISO (2016) protocol, to measure the lingering effects of a multitasking episode. Participants performed a pursuit-tracking task (Experiment 1) or drove in a high-fidelity driving simulator (Experiment 2) by itself or while concurrently

Table 9
Contrasts Between Baseline and Off-Task Pupil Diameter for the Hard Drive

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
Hard	*M*	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	5.10	0.17	5.42	.18	5.51	.21						
			Off-task −1's									
Bin 3			5.27	0.09			7.38	.001	0.27			
Bin 6			5.21	0.09			4.75	.001	0.17			
Bin 9			5.19	0.08			3.61	.002	0.13			
Bin 12			5.17	0.09			2.87	.027	0.10			
Bin 15			5.18	0.09			3.36	.005	0.12			
Bin 18			5.15	0.10			2.07	ns	0.08			
					Off-task −3's							
Bin 3					5.30	0.13	7.47	.001	0.32			
Bin 6					5.27	0.12	6.42	.001	0.27			
Bin 9					5.26	0.09	5.97	.001	0.25			
Bin 12					5.24	0.09	5.19	.001	0.22			
Bin 15					5.25	0.08	5.54	.001	0.23			
Bin 18					5.23	0.10	4.81	.001	0.20			
	5.10	0.17	5.40	.13	5.51	.14						
			Off-task −1's									
Bin 3			5.27	0.09			7.38	.001	0.27	−6.13	.001	−0.21
Bin 6			5.21	0.09			4.75	.001	0.17	−9.06	.001	−0.30
Bin 9			5.19	0.08			3.61	.002	0.13	−10.36	.001	−0.35
Bin 12			5.17	0.09			2.87	.027	0.10	−11.13	.001	−0.37
Bin 15			5.18	0.09			3.36	.005	0.12	−10.58	.001	−0.36
Bin 18			5.15	0.12			2.07	ns	0.08	−11.97	.001	−0.40
					Off-task −3's							
Bin 3					5.30	0.13	7.47	.001	0.32	−4.51	.001	−0.20
Bin 6					5.27	0.12	6.42	.001	0.27	−7.10	.001	−0.32
Bin 9					5.26	0.09	5.97	.001	0.25	−8.63	.001	−0.39
Bin 12					5.24	0.09	5.19	.001	0.22	−9.37	.001	−0.42
Bin 15					5.25	0.08	5.54	.001	0.23	−9.38	.001	−0.42
Bin 18					5.23	0.10	4.81	.001	0.20	−10.19	.001	−0.46

Figure 8

The Regions of Interest (ROIs) for the Forward Roadway and Instrument Panel

Note. ROI = regions of interest. The forward roadway ROI (outlined in red) encompassed 10.48% of the view in the scene camera and the instrument panel ROI (outlined in yellow) encompassed 5.86% of the view in the scene camera. Also visible in the image is the rear-view mirror (outlined in blue). The red, yellow, and blue outlines were not visible to the participant. See the online article for the color version of this figure.

performing an easy or difficult backwards counting task that diverted attention from the primary task of tracking or driving. The secondary counting task periodically started and stopped while participants performed the primary task. The DRT was used to probe the attentional demands throughout these on-task and off-task multitasking intervals.

As expected, DRT RT was higher and DRT Hit Rate was lower during the on-task interval compared to baseline levels with greater costs for the more demanding counting task (i.e., 3's vs. 1's). Surprisingly, these differences persisted long after the counting task had stopped. In both Experiments, DRT RT decreased as a negatively accelerated function of time with the biggest costs observed immediately after the counting task stopped. DRT RT remained above baseline levels throughout the 30-s off-task interval. DRT Hit Rate reached baseline levels after 12 s with the simple pursuit-tracking task in Experiment 1 and was below baseline levels throughout the 30-s off-task interval in the more complex driving task. Importantly, the pupillometry measures showed a similar pattern, establishing that the residual costs are obtained with both the DRT and physiological measures.

The Persistence of Distraction

The nature and extent of intermittent multitasking costs were largely hidden until the DRT was used to systematically probe for them. Estimates of lingering dual-task costs can be derived from the literature using the psychological refractory period (Pashler, 1984, 1994a) and task-switching paradigms (e.g., Rogers & Monsell, 1995). In both of these paradigms, the costs of performing a concurrent secondary task or switching from one task set to another tend to be relatively short-lived; on the order of a few seconds. The data presented in this report suggest that the dual-task literature may have underestimated the duration of multitasking costs by an order of magnitude.

Intuitively, one might think that the costs of intermittent multitasking would take some time to dissipate; however, the actual duration of these costs may be surprising. In fact, we found that the residual costs lasted longer than the duration of the secondary counting task. This pattern was also observed in the on-road testing involving voice-based interactions with the embedded infotainment system, which were often of shorter duration than the residual costs from these interactions (Strayer et al., 2016). Converging evidence from simple pursuit tracking, to simulated driving (see also Bowden et al., 2018), to driving an automobile on residential

Table 10

Contrasts Between Baseline, Off-Task and On-Task Time in Forward Roadway ROI (%) for the Easy Drive Q32

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	54.14	9.93	56.91	8.90	57.36	11.05						
			Off-task −1's									
Bin 3			48.89	13.07			−1.30	ns	−0.31	−2.94	.02	−0.20
Bin 6			51.10	12.65			−0.29	ns	−0.41	−1.92	ns	−0.13
Bin 9			49.89	10.70			−0.79	ns	−0.44	−2.43	ns	−0.17
Bin 12			49.62	9.76			−0.90	ns	−0.45	−2.55	ns	−0.17
Bin 15			49.74	8.62			−0.85	ns	−0.44	−2.50	ns	−0.17
Bin 18			47.99	12.33			−1.58	ns	−0.52	−3.23	.01	−0.22
					Off-task −3's							
Bin 3					52.54	8.00	0.40	ns	0.02	−2.692	.05	−0.19
Bin 6					50.24	7.77	−0.84	ns	−0.05	−2.288	ns	−0.16
Bin 9					56.80	9.52	2.73	.04	0.15	−0.315	ns	−0.02
Bin 12					54.88	8.20	1.59	ns	0.09	−1.387	ns	−0.10
Bin 15					54.69	8.93	1.54	ns	0.09	−0.763	ns	−0.05
Bin 18					52.01	7.76	0.10	ns	0.01	−1.808	ns	−0.13

Note. ROI = regions of interest.

Table 11
Contrasts Between Baseline, Off-Task and On-Task Time in Forward Roadway ROI (%) for the Hard Drive

Q33

	Baseline		On-task −1's		On-task −3's		Difference from baseline			difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	48.39	7.01	57.37	7.67	55.29	6.98						
			Off-task −1's									
Bin 3			50.84	8.79			0.758	ns	0.04	−3.36	.01	−0.17
Bin 6			49.81	8.41			0.190	ns	0.01	−3.68	.002	−0.19
Bin 9			49.91	9.52			0.246	ns	0.01	−3.82	.001	−0.20
Bin 12			48.61	7.74			−0.469	ns	−0.02	−4.30	.001	−0.22
Bin 15			49.66	8.08			0.105	ns	0.01	−4.12	.001	−0.21
Bin 18			48.42	8.17			−0.575	ns	−0.03	−4.53	.001	−0.23
					Off-task −3's							
Bin 3					50.40	6.94	0.56	ns	0.02	−3.57	.003	0.02
Bin 6					53.74	6.48	2.58	ns	−0.05	−1.50	.001	−0.05
Bin 9					50.42	7.66	0.58	ns	0.15	−3.65	.002	0.15
Bin 12					52.76	9.14	1.99	ns	0.09	−2.12	.001	0.09
Bin 15					54.19	6.60	2.85	.03	0.08	−1.18	.02	0.08
Bin 18					54.62	7.87	3.11	.01	0.01	−0.94	.001	0.01

Note. ROI = regions of interest.

streets document the long-lasting costs of intermittent multitasking (see also Jenness et al., 2015). Moreover, these costs were obtained with a standardized backwards counting task and, as shown by Strayer et al. (2016), with secondary tasks involving the selection of music, placing a phone call, or using voice commands to send a text message while driving a motor vehicle. These residual costs also tended to be greater when the secondary task was more complex.

One possible explanation for the residual costs is that drivers were attempting to reacquire situation awareness that was lost while multitasking (e.g., Strayer & Fisher, 2016). However, the eye-tracking measures did not show any systematic evidence of enhanced visual scanning in the off-task interval that would suggest an oversampling of the primary task (e.g., scanning the periphery and/or looking at the instrument cluster). Moreover, a similar pattern of residual costs was observed in the simple pursuit-tracking task, where it is not obvious what situation awareness was to be regained in the off-task interval. That is, a more benign interpretation of the residual costs—reacquiring situation awareness that was lost while multitasking—does not appear to be supported by these data. It is also noteworthy that these impairments are cognitive in nature because there was no visual information for participants to look at while performing the secondary counting tasks (but see Bowden et al., 2018).

That the residual costs were observed with both the DRT measures (RT and Hit Rate) and with pupil diameter is also important because it indicates that the residual costs are not an artifact of the DRT measurement protocol. Moreover, the fact that the residual costs are greater with more demanding secondary tasks rules out simple sensory and motoric interpretations of the effect. Rather, the impairments appear consistent with the proactive interference hypothesis. From the perspective of the threaded cognition (Salvucci & Taatgen, 2008), instead of the threads being "politely released" when no longer needed, the threads appear to be "sticky," gumming up the works until the information is slowly purged from working memory. By contrast, the time to load new procedures into working memory and to observe the multitasking costs is much shorter. From the perspective of threaded cognition (Salvucci & Taatgen, 2008), these new multitasking threads are "greedy," requesting processing resources as soon as possible.

The DRT protocol involves presenting a simple stimulus and requiring the participant to make a simple button press (a simple RT task). How might threaded cognition account for the sensitivity of the DRT to primary and secondary-task demand? One possibility would be that a DRT thread gets added to the goal buffer. For example,

IF the goal buffer contains a triple task (e.g., the driving task, a secondary task, and the DRT task)

THEN add the goal to perform the driving task

and add the goal to perform the secondary task

and add the goal to perform the DRT task.

…

If the goal buffer contains the DRT task

and the DRT stimulus has been presented

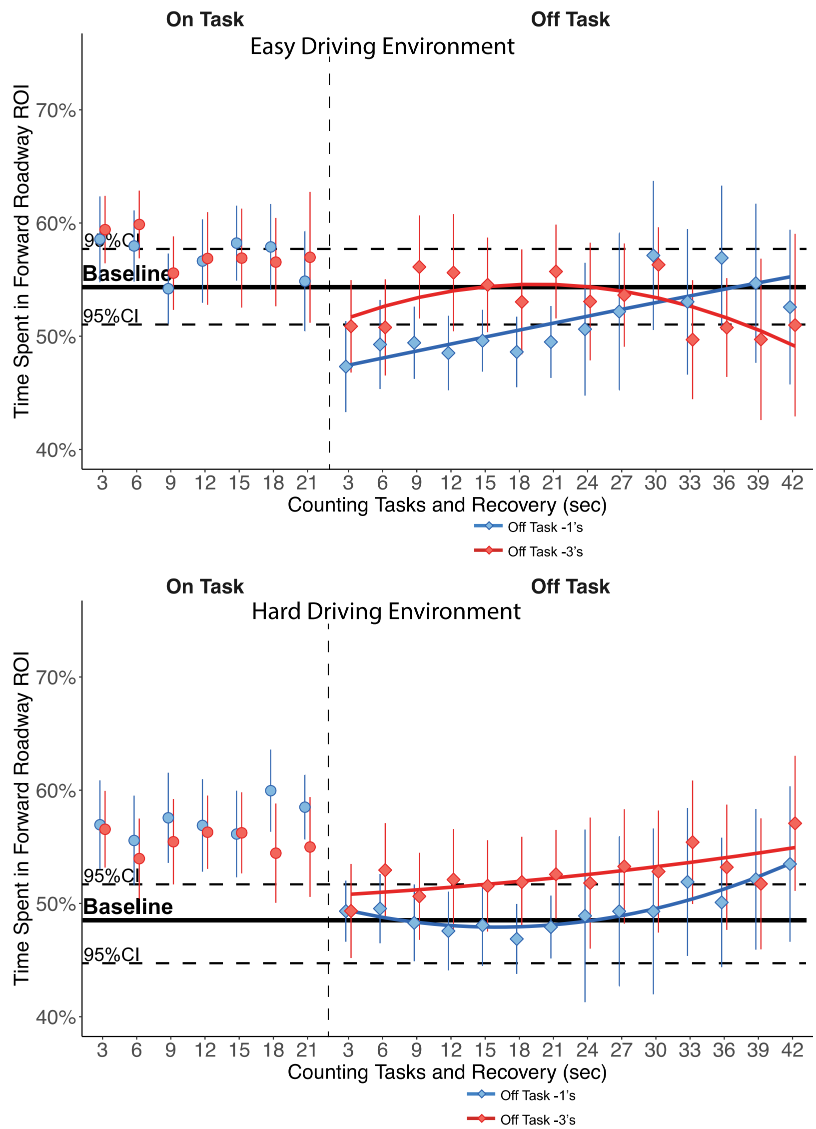
Then issue the DRT motor response.

The exclusive-use rule of threaded cognition would cause the DRT thread to wait its turn in the goal buffer until other threads had been completed. Consequently, the more demanding the driving task and/or the more demanding the secondary task, the longer the latency of the DRT response. Note that because of the simplicity of the DRT task, the DRT thread should take few processing resources and would therefore produce little interference with other concurrent tasks.

From an evidence accumulation modeling perspective (e.g., Brown & Heathcote, 2008), the slow return to single-task baseline levels may be indicative of a gradual decrease in response thresholds

Figure 9

Q34 *Percentage of Time Spent in the Forward Roadway Regions of Interest (ROI) in Experiment 2*



ONLINE COLOR

Note. The gray vertical dashed lines indicate the point at which participants ceased counting. In the Easy driving environment (top panel), the solid blue and red lines represent the best-fitting power functions of an off-task percentage of time for 1's (R^2 = .79) and 3's (R^2 = .63), respectively. In the Hard driving environment (bottom panel), the solid blue and red lines represent the best-fitting power functions of the off-task percentage of time for 1's (R^2 = .88) and 3's (R^2 = .84), respectively. See the online article for the color version of this figure.

that were elevated when the participant was actively multitasking Q39 (e.g., Castro et al., 2019; Howard et al., 2018; Tillman et al., 2017). A slow decrease in response thresholds has also been documented in other contexts (e.g., Strayer & Kramer, 1994) and is consistent with the proactive interference interpretation.

Finally, given that residual costs have been observed in primary-task measures such as lane keeping and driving speed (e.g., Bowden et al., 2018; Jenness et al., 2015), secondary-task DRT measures (see also Bowden, et al., 2018; Strayer et al., 2016, Strayer, Biondi, et al., 2017; Strayer, Cooper, et al., 2017) Q40 and physiological measures (e.g., pupil diameter), it is likely that motorists on the roadway are impaired after completing a multitasking episode. For example, one potential consequence of the residual multitasking costs may occur at intersections controlled by traffic lights. The last few years have seen significant year-over-year increases in pedestrian and bicycle fatalities

Table 12
Contrasts Between Baseline, Off-Task and On-Task Time in the Instrument Panel ROI (%) for the Easy Drive Q36

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	5.73	4.03	2.29	4.48	1.51	4.93						
			Off-task −1's									
Bin 3			5.45	5.93			−0.03	ns	−0.00	2.85	.03	0.43
Bin 6			5.23	4.11			−0.27	ns	−0.03	2.66	.05	0.40
Bin 9			5.68	4.81			0.20	ns	0.03	3.06	.01	0.47
Bin 12			6.62	5.67			1.09	ns	0.14	3.89	.001	0.59
Bin 15			6.29	4.68			0.78	ns	0.10	3.60	.002	0.55
Bin 18			6.34	6.21			0.89	ns	0.12	3.64	.002	0.56
					Off-task −3's							
Bin 3					4.79	4.23	−0.861	ns	0.02	4.66	.001	0.73
Bin 6					5.05	3.64	−0.660	ns	−0.05	2.98	.02	0.47
Bin 9					3.53	3.42	−2.418	ns	0.15	2.04	ns	0.32
Bin 12					4.44	4.71	−1.319	ns	0.09	3.12	.01	0.49
Bin 15					4.10	3.36	−1.739	ns	0.08	2.52	ns	0.40
Bin 18					4.15	3.31	−1.652	ns	0.01	1.86	ns	0.29

Note. ROI = regions of interest.

compared to other crash types. In fact, in 2018 there was a 2.4% *decrease* in overall traffic fatalities from the prior year, yet in that same interval there was a 3% *increase* in pedestrian fatalities and a 6% *increase* in bicycle fatalities (National Highway Traffic Safety Administration [NHTSA], 2018). Many drivers use their mobile devices to send or read text messages when stopped at a red light (DOT, 2019). When the light turns green these multitasking drivers may proceed thru the intersection without fully scanning for potential hazards. Converging sources of evidence suggest that when these motorists stop texting, they may still be impaired by the lingering distraction caused by multitasking.

Conclusion

Multitasking is ubiquitous in modern society. While the costs incurred while actively performing functionally independent tasks are reasonably well understood, the hidden costs incurred after intermittent multitasking have not been adequately explored. We found that the costs of intermittent multitasking persist much longer than predicted based on the dual-task literature. Residual costs lasting a half-minute or more were observed when a cognitively demanding secondary task was paired with (a) simple pursuit tracking, (b) driving a high-fidelity simulator, and, in earlier research, and (c) operating a motor vehicle on residential streets. We suggest that it

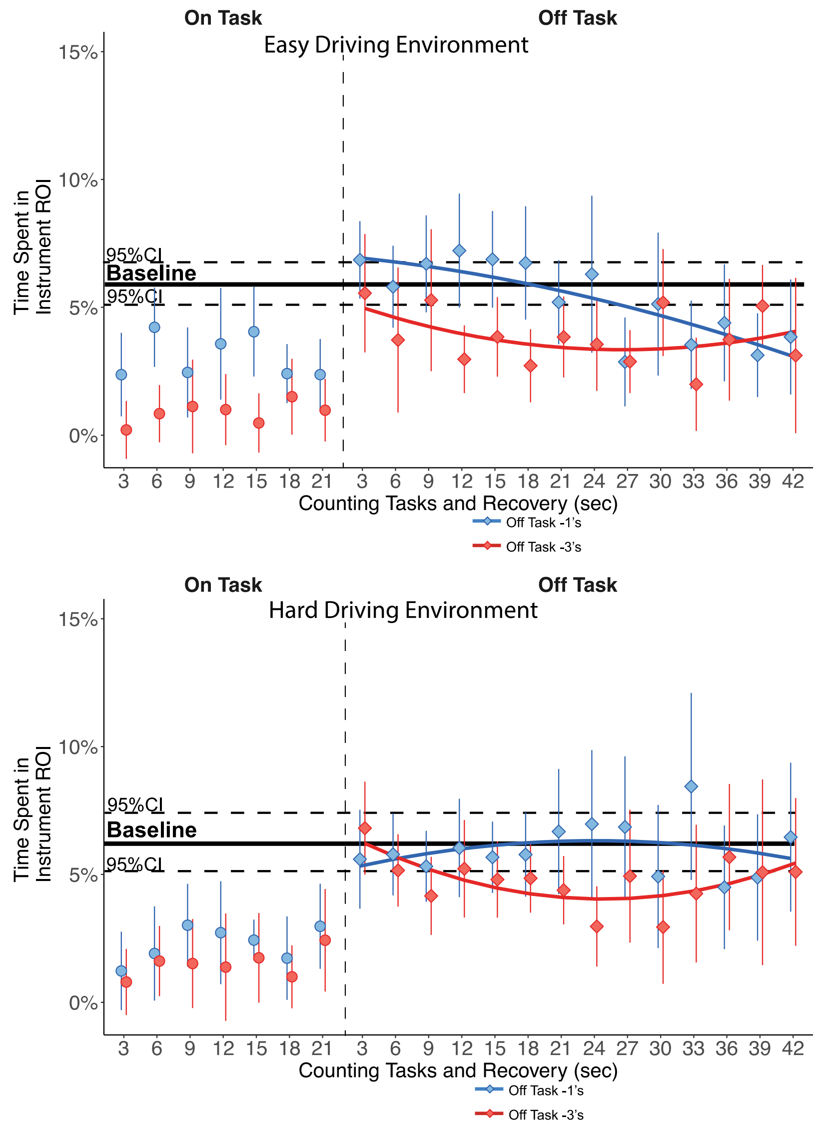
Table 13
Contrasts Between Baseline, Off-Task and On-Task Time in the Instrument Panel ROI (%) for the Hard Drive Q37

	Baseline		On-task −1's		On-task −3's		Difference from baseline			Difference from on-task		
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*(46)	*p*	*d*	*t*(46)	*p*	*d*
	6.41	4.21	3.06	5.15	0.88	4.90						
			Off-task −1's									
Bin 3			7.47	6.58			0.83	ns	0.11	3.728	.001	0.54
Bin 6			6.46	5.09			−0.05	ns	−0.01	2.644	.05	0.38
Bin 9			6.74	4.81			0.20	ns	0.03	3.069	.01	0.45
Bin 12			8.24	6.05			1.49	ns	0.20	4.102	.001	0.60
Bin 15			6.92	4.92			0.36	ns	0.05	3.370	.01	0.49
Bin 18			6.82	4.94			0.27	ns	0.04	3.157	.01	0.46
					Off-task −3's							
Bin 3					6.08	7.30	−0.36	ns	−0.06	4.54	.001	0.74
Bin 6					3.84	6.44	−2.26	ns	−0.35	2.54	ns	0.41
Bin 9					5.44	6.42	−0.91	ns	−0.14	3.98	.001	0.65
Bin 12					3.33	4.07	−2.69	.05	−0.41	2.15	ns	0.35
Bin 15					3.53	2.82	−2.52	ns	−0.39	2.23	ns	0.36
Bin 18					2.43	3.42	−3.45	ns	−0.53	1.25	ns	0.20

Note. ROI = regions of interest.

Figure 10

Percentage of Time Spent in the Instrument Panel Regions of Interest (ROI) in Experiment 2



ONLINE COLOR

Note. The gray vertical dashed lines indicate the point at which participants ceased counting. In the Easy driving environment (top panel), the solid blue and red lines represent the best-fitting power functions of the off-task percentage of time for 1's (R^2 = .61) and 3's (R^2 = .57), respectively. In the Hard driving environment (bottom panel), the solid blue and red lines represent the best-fitting power functions of the off-task percentage of time for 1's (R^2 = .63) and 3's (R^2 = .66), respectively. See the online article for the color version of this figure.

takes time to unburden working memory from multitasking. Whereas loading new procedures into working memory occurs fairly quickly, the purging of this information takes considerably longer.

References

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with scholastic aptitude test scores. *Science*, *205*, 1289–1292. https://doi.org/10.1126/science.472746

Q41

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.005

tBeatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Bernston (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142–162). Cambridge University Press.

Bergen, B., Medeiros-Ward, N., Wheeler, K., Drews, F., & Strayer, D. L. (2013). The crosstalk hypothesis: Language interferes with driving because of modality-specific mental simulation. *Journal of Experimental Psychology: General*, *142*, 119–130. https://doi.org/10.1037/a0028428

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178. https://doi.org/10.1016/j.cogpsych.2007.12.002

Carsten, O. M. J., Merat, N., Janssen, W. H., Johansson, E., Fowkes, M., & Brookhuis, K. A. (2005). *HASTE final report*. Institute for Transportation Studies, University of Leeds.

Castro, S. C., Strayer, D. L., Matzke, D., & Heathcote, A. (2019). Cognitive workload measurement and modeling under divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 826–839. https://doi.org/10.1037/xhp0000638

Cooper, J. M., Castro, S. C., & Strayer, D. L. (2016, September). Extending the Detection Response Task to simultaneously measure cognitive and visual task demands. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 60, No. 1, pp. 1962–1966). Sage Publications. https://doi.org/10.1177/1541931213601447

Q42 Cooper, J. M., Medeiros-Ward, N., & Strayer, D. L. (2013). The impact of eye movements and cognitive workload on lateral position variability in driving. *Human Factors*, *55*, 1001–1014. https://doi.org/10.1177/0018720813480177

Department of Transportation. (2019). *Driver electronic device use in 2018* (DOT HS 818 818). Traffic Safety Facts: U.S. Department of Transportation National Highway Traffic Safety Administration.

Durso, F., Rawson, K., & Girotto, S. (2007). Comprehension and situation awareness. In F. T. Durso, R. Nickerson, S. T. Dumais, S. Lewandowsky, & T. Perfect (Eds.), *The handbook of applied cognition* (2nd ed.). Wiley. https://doi.org/10.1002/9780470713181.ch7

Endsley, M. R. (1995). Towards a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Endsley, M. R. (2015). Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making*, *9*, 4–32. https://doi.org/10.1177/1555343415572631

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, *37*(2), 381–394. https://doi.org/10.1518/001872095779064555

Q43 Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23. https://doi.org/10.1111/1467-8721.00160

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Horrey, W. J., Wickens, C. D., & Consalus, K. P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied*, *12*(2), 67–78. https://doi.org/10.1037/1076-898X.12.2.67

Howard, Z. L., Evans, N. J., Innes, R. J., Brown, S. D., & Eidels, A. (2020). How is multi-tasking different from increased difficulty? *Psychonomic Bulletin & Review*, *27*, 937–951. https://doi.org/10.3758/s13423-020-01741-8

Hughes, J. (2017). Reghelper: Helper functions for regression analysis (R package Version 0.3. 4). https://CRAN.R-project.org/package=reghelper

Q45 Q44 ISO DIS 17488. (2016). Road vehicles—Transport information and control systems—Detection-Response Task (DRT) for assessing attentional effects of cognitive load in driving. ISO TC 22/SC39/WG8.

Q46 James, W. (1890). *The principles of psychology*. Harvard University Press.

Jenness, J. W., Baldwin, C., Chrysler, S., Lee J. D., et al. (2015). Connected vehicle DVI design research and distraction assessment. *HFCV Phase 2 NHTSA DTNH22-11-D-00237 Task Order 0001*.

Q48 Q47

Joshi, S., & Gold, J. I. (2020). Pupil size as a window on neural substrates of cognition. *Trends in Cognitive Sciences*, *24*(6), 466–480. https://doi.org/10.1016/j.tics.2020.03.005

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*, 1583–1585. https://doi.org/10.1126/science.154.3756.1583

Kass, S. J., Cole, K. S., & Stanny, C. J. (2007). Effects of distraction and experience on situation awareness and simulated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *10*, 321–329. https://doi.org/10.1016/j.trf.2006.12.002

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking-An integrative review of dual-task and task-switching research. *Psychological Bulletin*, *144*, 557–583. https://doi.org/10.1037/bul0000144

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13 Q49

LoTemplio, S., Silcox, J., Federmeier, K. D., & Payne, B. R. (2021). Inter- and intra-individual coupling between pupillary, electrophysiological, and behavioral responses in a visual oddball task. *Psychophysiology*, *58*, Article e13758. https://doi.org/10.1111/psyp.13758 Q50

Medeiros-Ward, N., Cooper, J. M., & Strayer, D. L. (2014). Hierarchical control and driving. *Journal of Experimental Psychology: General*, *143*, 953–958. https://doi.org/10.1037/a0035097

National Highway Traffic Safety Administration. (2018). *Traffic deaths decreased in 2018, but still 36,560 people died.*https://www.nhtsa.gov/traffic-deaths-2018 Q51

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*, 214–255. https://doi.org/10.1037/0033-295X.86.3.214

Navon, D., & Miller, J. (1987). Role of outcome conflict in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 435–448. https://doi.org/10.1037/0096-1523.13.3.435

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*, 44–64. https://doi.org/10.1016/0010-0285(75)90004-3

Normark, C. J., Tretten, P., & Gärling, A. (2009, June). Do redundant head-up and head-down display configurations cause distractions. In *Proceedings of the 5th international driving symposium on human factors in driver assessment and design* (pp. 398–404). Q52

Palada, H., Neal, A., Strayer, D., Ballard, T., & Heathcote, A. (2019). Using response time modeling to understand the sources of dual-task interference in a dynamic environment. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 1331–1345. https://doi.org/10.1037/xhp0000672

Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 358–377. https://doi.org/10.1037/0096-1523.10.3.358

Pashler, H. (1994a). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*, 220–244. https://doi.org/10.1037/0033-2909.116.2.220 Q53

Pashler, H. (1994b). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 330–342. https://doi.org/10.1037/0096-1523.20.2.330 Q54

Pashler, H. (2000). Task switching and multitask performance. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 277–309). The MIT Press. Q55

Pellecchia, G. L., Shockley, K., & Turvey, M. T. (2005). Concurrent cognitive task modulates coordination dynamics. *Cognitive Science*, *29*(4), 531–557. https://doi.org/10.1207/s15516709cog0000_12

Poitras, M., Péléja, L., Lavertu, G., Langlois, A., Boulerice, K., Berthelot, P., Vincent-Lamarre, P., Beaulieu, S., Bournival, V., Brault, L., Charlebois, J., Galloway, E., Gauthier, A., Gibeau, R., & Giroux, N. (2020). A replication of Waugh and Norman (1965) primary memory study. *The Quantitative Methods for Psychology*, *16*, r1–r7. https://doi.org/10.20982/tqmp.16.2.r001

Q56 Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, *78*, 391–408. https://doi.org/10.1037/h0031333

Posner, M. I., & Keele, S. W. (1967). Decay of visual information from a single letter. *Science*, *158*, 137–139. https://doi.org/10.1126/science.158.3797.137

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org/

Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident; Analysis and Prevention*, *43*, 1771–1781. https://doi.org/10.1016/j.aap.2011.04.008

Reimer, B. (2009). Impact of cognitive task complexity on drivers' visual tunneling. *Transportation Research Record: Journal of the Transportation Research Board*, *2138*(1), 13–19.

Q57 Rogers, R. D., & Monsell, S. D. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review*, *115*, 101–130. https://doi.org/10.1037/0033-295X.115.1.101

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who Multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PLOS ONE*, *8*(1), Article e54402. https://doi.org/10.1371/journal.pone.0054402

Silcox, J. W., & Payne, B. R. (2021). The costs (and benefits) of effortful listening on context processing: A simultaneous electrophysiological, pupillometry, and behavioral study [Preprint]. *PsyArXiv*. https://doi.org/10.31234/osf.io/f8pzg

Strayer, D. L., Biondi, F., & Cooper, J. M. (2017). Dynamic workload fluctuations in driver/non-driver conversational dyads. In D. V. McGehee, J. D. Lee, & M. Rizzo (Eds.), *Driving assessment 2017: International symposium on human factors in driver assessment, training, and vehicle design* (pp. 362–367). Public Policy Center, University of Iowa.

Strayer, D. L., Cooper, J. M., Goethe, R. M., McCarty, M. M., Getty, D. J., & Biondi, F. (2019). Assessing the visual and cognitive demands of in-vehicle information systems. *Cognitive Research: Principles and Implications*, *4*, Article 18. https://doi.org/10.1186/s41235-019-0166-3

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J. R., & Hopman, R. J. (2016). Talking to your car can drive you to distraction. *Cognitive Research: Principles and Implications*, *1*, Article 16. https://doi.org/10.1186/s41235-016-0018-3

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J. R., & Hopman, R. J. (2017). The smartphone and the driver's cognitive workload: A comparison of Apple, Google, and Microsoft's intelligent personal assistants. *Canadian Journal of Experimental Psychology*, *71*, 93–110. https://doi.org/10.1037/cep0000104

Strayer, D. L., & Fisher, D. L. (2016). SPIDER: A framework for understanding driver distraction. *Human Factors*, *58*(1), 5–12. https://doi.org/10.1177/0018720815619074

Strayer, D. L., & Kramer, A. F. (1994). Strategies and automaticity II: Dynamic aspects of strategy adjustment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 342–365. https://doi.org/10.1037/0278-7393.20.2.342

Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., & Biondi, F. (2015). Assessing cognitive distraction in the automobile. *Human Factors*, *57*, 1300–1324. https://doi.org/10.1177/0018720815575149

Strayer, D. L., Watson, J. M., & Drews, F. A. (2011). Cognitive distraction while multitasking in the automobile. *Psychology of Learning and Motivation*, *54*, 29–58. https://doi.org/10.1016/B978-0-12-385527-5.00002-4

Tillman, G., Strayer, D., Eidels, A., & Heathcote, A. (2017). Modeling cognitive load effects of conversation between a passenger and driver. *Attention, Perception & Psychophysics*, *79*, 1795–1803. https://doi.org/10.3758/s13414-017-1337-2

von der Malsburg, T. (2015). Saccades: Detection of fixations in eye-tracking data (R package Version 0.1-1). https://CRAN.R-project.org/package=saccades Q58

Wang, Y., Reimer, B., Dobres, J., & Mehler, B. (2014). The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation Research Part F: Traffic Psychology and Behaviour*, *26*, 227–237. https://doi.org/10.1016/j.trf.2014.08.003

Watson, J. M., & Strayer, D. L. (2010). Supertaskers: Profiles in extraordinary multitasking ability. *Psychonomic Bulletin & Review*, *17*, 479–485. https://doi.org/10.3758/PBR.17.4.479

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance* (Vol. 8, pp. 239–257). Lawrence Erlbaum.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 63–101). Academic Press.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*, 449–455. https://doi.org/10.1518/001872008X288394

World Health Organization. (2011). *Mobile phone use: A growing problem of driver distraction*. http://www.who.int/violence_injury_prevention/publications/road_traffic/en/index.html

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology*, *51*, 277–284. https://doi.org/10.1111/psyp.12151

Received November 6, 2020
Revision received April 27, 2021
Accepted June 30, 2021 ■

Queries

Q1. Please check and confirm the given names and surnames of the author(s) are identified correctly and amend, if necessary. Names currently identified as surnames are highlighted.

Q2. Kindly check and confirm whether the department name is correctly identified for the affiliation "1".

Q3. Please provide department details for the affiliation "2".

Q4. Please provide department/city/state/country name for the affiliations "3 & 4".

Q5. Kindly check and confirm whether the keywords are correctly identified.

Q6. 'Castro, Heathcote, Cooper & Strayer, under review' is cited but the corresponding reference not provided.

Q7. Please provide expansion for abbreviation "RT" used in the article.

Q8. The citation "Palada, et al., (2017)" has been changed to "Palada et al. (2019)" to match the author name/date in the reference list. Please check here and in subsequent occurrences, and correct if necessary.

Q9. The citation "Strayer et al., 2017" has been changed to "Strayer, Biondi, et al., 2017; Strayer, Cooper, et al., 2017". Kindly check and remove the incorrect citation from this occurrence.

Q10. 'Bowden et al. (2018' is cited but the corresponding reference not provided.

Q11. The citation "Strayer et al., 2017" has been changed to "Strayer, Biondi, et al., 2017; Strayer, Cooper, et al., 2017". Kindly check and remove the incorrect citation from this occurrence.

Q12. 'Fisher and Strayer (2014)' is cited but the corresponding reference not provided.

Q13. Please provide expansion for "RMSEA, ANOVA" used in the article.

Q14. Please provide expansion for "DRT RT" used in the article.

Q15. Please provide expansion for the abbreviations "DRT RT & RT" used in the Figures 2 and 5.

Q16. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 1.

Q17. Please provide expansion for "DRT RT" used in the Tables 1, 4 & 5.

Q18. Kindly check and confirm the edit made in the Figures 3 & 6 foot notes.

Q19. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 2.

Q20. Kindly check and confirm whether the Tables 2, 6 & 7 foot note is correctly identified.

Q21. The citation "Strayer et al., 2017" has been changed to "Strayer, Biondi, et al., 2017; Strayer, Cooper, et al., 2017". Kindly check and remove the incorrect citation from this occurrence.

Q22. 'Drews et al., 2008' is cited but the corresponding reference not provided.

Q23. Trademark symbol has been remove here. Kindly check and confirm.

Q24. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 3.

Q25. Please provide a definition for the significance of [italics] in the Table 3.

Q26. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 4.

Q27. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 5.

Q28. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 6.

Q29. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 7.

Q30. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 8.

Q31. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 10.

Q32. Kindly check and confirm whether the Table foot note is correctly identified. For Tables 10–13.

Q33. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 11.

Q34. Kindly check and confirm the edit made in the Figures 9 & 10 caption.

Q35. Please check and confirm the inserted citation of Figure 9.

Q36. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 12.

Q37. As per APA guidelines, all table columns should have a header. Please provide headers for column 1 for Table 13.

Q38. Please check and confirm the inserted citation of Figure 10.

Q39. 'Howard et al., 2018' is cited but the corresponding reference not provided.

Q40. The citation "Strayer et al., 2017" has been changed to "Strayer, Biondi, et al., 2017; Strayer, Cooper, et al., 2017". Kindly check and remove the incorrect citation from this occurrence.

Q41. Please provide a missing issue number for all the journal type references.

Q42. 'Cooper et al., 2016' is not cited in text. Please add citation or delete from list.

Q43. 'Endsley and Kiris, 1995' is not cited in text. Please add citation or delete from list.

Q44. Please provide publisher name for the reference (Ref. "Hughes, 2017").

Q45. 'Hughes, 2017' is not cited in text. Please add citation or delete from list.

Q46. Please provide complete details for the reference (Ref. "ISO DIS 17488, 2016").

Q47. Please provide complete details for the reference (Ref. "Jenness et al., 2015").

Q48. As per journal style "et al." Is not allowed in the reference list. Kindly check and update the complete author names for the reference (Ref. "Jenness et al., 2015").

Q49. 'Kuznetsova et al., 2017' is not cited in text. Please add citation or delete from list.

Q50. Year has been changed from "2020" to "2021" in both reference list and citation (Ref. "LoTemplio et al., 2021"). Kindly check and confirm.

Q51. Kindly check and confirm whether the webtitle is correctly identified (Ref. "National Highway Traffic Safety Administration, 2018").

Q52. Please provide either editors name/publisher name or Conference location for the reference (Ref. "Normark et al., 2009").

Q53. Year has been changed from "1994" to "1994a" in both reference list and citations (Ref. "Pashler, 1994a"). Kindly check and confirm.

Q54. 'Pashler, 1994b' is not cited in text. Please add citation or delete from list.

Q55. 'Pashler, 2000' is not cited in text. Please add citation or delete from list.

Q56. Author names (Vincent-Lamarre, P., Beaulieu, S., Bournival, V., Brault, L., Charlebois, J., Galloway, E., Gauthier, A., Gibeau, R., & Giroux, N.) has been added for the reference (Ref. "Poitras et al., 2020'). Kindly check and confirm.

Q57. Please provide DOI number for the reference (Ref. "Reimer, 2009").

Q58. Please provide publisher name for the reference (Ref. "von der Malsburg, 2015").