

Examining the effects of goal-setting, feedback, and incentives on sustained attention

Matthew K. Robison¹, Nash Unsworth², & Gene A. Brewer³

¹ University of Texas at Arlington

² University of Oregon

³ Arizona State University

Abstract

Across four experiments we examined the effects of goal-setting, feedback, and incentivizing manipulations on sustained attention. In addition to measuring task performance, we measured subjective attentional states and momentary feelings of motivation and alertness. Experiment 1 compared two specific goal conditions - one difficult and one easy - to a standard set of instructions. The specific goal conditions both reduced RTs and attenuated the vigilance decrement but did not impact task engagement (motivation or task-unrelated thoughts). Experiment 2 manipulated both goal-setting and feedback across conditions. The combination of a specific goal and feedback had strong effects on both task performance and task engagement. Additionally, feedback increased task engagement (higher motivation and fewer task-unrelated thoughts) regardless of whether or not it was paired with a specific goal. Experiment 3 examined the effect of pairing goals with a reward. Participants in one reward condition (time-based incentive) reported higher motivation, but did not show better task performance. Offering a cash incentive to meet a goal did not have an effect on any dependent variables. Finally, in an effort to examine whether more moderately-difficult goals might lead to optimal performance, Experiment 4 examined a broader range of goals. However we did not see an effect of a moderately-difficult goal on any of the dependent variables. Although some of the experimental manipulations were effective in mitigating the vigilance decrement, none eliminated it. We discuss the theoretical implications of the results with regard to goal-setting theory and theories of vigilance.

Keywords: sustained attention; vigilance; motivation; goal-setting; mind-wandering

Word count: 10989

Public Significance Statement

Almost everyone can recognize that sustaining attention, even to important tasks like driving, reading, and listening in meetings, can be quite difficult. An active debate exists as to why this is the case. Some groundbreaking studies in the middle of the 20th century, combined with some recent experimentation, have shown promising avenues for improving our ability to sustain attention. These experiments have called into question longstanding theoretical arguments for why sustaining attention is so difficult. In the present study, we seek to extend this work to test theoretical stances regarding sustained attention, and perhaps find practically useful ways of making sustained attention situations more engaging.

Matthew K. Robison, Department of Psychology, University of Texas at Arlington; Nash Unsworth, Department of Psychology, University of Oregon; Gene A. Brewer, Department of Psychology, Arizona State University. Portions of these data were collected as part of a dissertation by M. K. R, in partial fulfillment of the requirements for degree of Doctor of Philosophy at the University of Oregon.

All authors were funded by National Science Foundation grant 1632327 during data collection and writing of this article.

Correspondence concerning this article should be addressed to Matthew K. Robison, 501 Nedderman Drive, Box 19528, Arlington, TX 76019. E-mail: matthew.robison@uta.edu

Sustaining one's attention is notoriously difficult. Psychologists have been interested in this fact for over a century (Bills, 1931b, 1931a; Thorndike, 1912). But research into sustained attention increased in the post-World War II era after Mackworth designed his Clock Test to examine prolonged performance on important but monotonous tasks (e.g., radar monitoring). Mackworth found that the longer someone spent doing his Clock Test, the worse their performance tended to be (Mackworth, 1948). Since that seminal study, the same effect, which came to be known as the *vigilance decrement*, has been replicated in a myriad of tasks. Cognitive psychologists have spent decades trying to drill down precisely why the vigilance decrement occurs, and why sustaining attention is so hard for most people (see Thomson, Besner, & Smilek, 2015; Esterman & Rothlein, 2019 for recent reviews).

One crucial question remains under debate: what causes the vigilance decrement? Several theories have been proposed. One major theory - resource theory - argues that vigilance situations require cognitive resources because they are demanding and stressful (Warm, Parasuraman, & Matthews, 2008). Over time, these resources become depleted, causing decrements to task performance. Mindlessness theory argues that vigilance situations are actually cognitively underwhelming, and the monotony causes people to intentionally disengage from the task as a means of escaping the monotony (Manly, Robertson, Galloway, & Hawkins, 1999; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997). Thomson et al. (2015) argue that the availability of cognitive resources does not change over time, but rather that people reallocate resources away from externally-directed attention (i.e., the task) and toward internally-directed thoughts (i.e., mind-wandering). Finally, some recent theories argue that there are costs to engaging in sustained attention, and gradually these costs start to outweigh the benefits of continued effort (Kurzban, Duckworth, Kable, & Myers, 2013). We address how our findings fit with predictions made by such theories in the General Discussion. Here, our goal here was to determine whether various exogenous factors could affect the degree to which people engage with a sustained attention task by taking both objective (i.e., task performance, vigilance decrement) and subjective (i.e., motivation, alertness, task-unrelated thoughts) measures into account.

More recently, there has been a surge of research on mind-wandering - occasional shifts of attention away from an ongoing task toward irrelevant internal streams of thought. Mind-wandering is particularly common during tasks that require sustained attention (McVay & Kane, 2009, 2012; Seli, Cheyne, & Smilek, 2013; Smallwood et al., 2004; Unsworth & Robison, 2016). Several studies have shown that mind-wandering and task performance can be affected by manipulating aspects of tasks. For example, Seli, Schacter, Risko, and Smilek (2019) told participants in one condition that they would be allowed to leave a 1-hour experiment after 30 min, contingent on their performance. A control condition did not receive such instructions. Compared to the control condition, participants with the motivational instruction reported higher motivation and less mind-wandering and also achieved better task performance. Thus, the subjective experience of mind-wandering is clearly manipulable, and it is accompanied by changes in motivation and task performance. As other examples, researchers have shown that mind-wandering can be manipulated by changing the difficulty of a task (e.g., Rummel & Boywitt, 2014; Feng, D'Mello, & Graesser, 2013; Thomson, Besner, & Smilek, 2013; Xu & Metcalfe, 2016). In the present study, we use thought probes to measure mind-wandering and other task-unrelated thoughts during the course

of a sustained attention task, and we leverage this prior work to examine whether various manipulations can affect the vigilance decrement. If so, we can start to test predictions made by competing theories of vigilance.

In a separate yet related line of research, industrial/organizational psychologists have been researching means of improving and maximizing task performance in both basic and applied settings. One prominent approach is goal-setting (Locke & Latham, 1990, 2002, 2019). According to goal-setting theory, one generally effective strategy for maximizing task performance - whether it be a simple task like solving math problems (Locke & Bryan, 1969) or a complex task like logging forests (Latham & Kinne, 1974) - is to set specific, difficult goals. Locke and Latham (2002) argue that goals affect performance in four ways: 1) they direct attention and effort toward a task, 2) they increase the level of effort toward that task, 3) they prolong effort, and 4) they motivate the self-development of strategies to increase performance. The third mechanism is particularly salient when considering the vigilance decrement. That is, Locke and Latham (2002) argue that setting goals should stabilize effort across a prolonged period, as effort and task performance typically decline as a function of time-on-task. Whereas vigilance researchers are specifically interested in the role of attentiveness to a cognitively demanding task, Locke and Latham note that this effect occurs across a wide variety of occupational settings, and that goals may be a way to mitigate this effect. As part of the theory, Locke and Latham also point to both mediators and moderators of the effect of goals on performance. As mediators, Locke and Latham identify choice, persistence, effort, and generation of strategies. As moderators, Locke and Latham identify feedback, goal commitment, and ability. In the present study, we operationalized goals by setting specific criteria for average RTs, we operationalized feedback by providing/removing trial- and block-level performance feedback, and we operationalized motivation and alertness by allowing participants to self-report their current levels of each. By doing so we created a set of independent and dependent variables that were tethered to the specific task employed here.

In the present study we applied goal-setting theory to sustained attention. If goals can improve sustained attention, goal-setting manipulations could open doors for using these manipulations to test theories regarding the vigilance decrement. Prior research has attempted to improve sustained attention with various other manipulations. For example, Brewer, Lau, Wingert, Ball, and Blais (2017) had participants complete two runs of the psychomotor vigilance task in dyads. After the first run, they offered a \$10 gift card to the participant who emitted fewer response times longer than 500 ms. Although they found a general speeding of responses during the second, motivated run, this incentive did not eliminate or attenuate the vigilance decrement. As another example, Hopstaken et al. (2015a) had participants complete five blocks of a sustained attention task for 30 minutes. After the fifth block, they informed participants that the time remaining in the experiment would depend on their performance on the subsequent block. Participants were told the task could end in as few as five minutes or take as long as an additional 30 minutes, depending on their performance. They found a significant improvement in performance and physiological measures of engagement following this instruction. These data suggest that people can be sufficiently encouraged to re-engage with a sustained attention task, even after experiencing typical decrements across time. In other examples, Massar et al. (2016,

2019) demonstrated improvements to performance in the psychomotor vigilance task when participants were monetarily rewarded on a trial-by-trial basis for fast reaction times, and Esterman, Reagan, Liu, Turner, and DeGutis (2014) showed improvements on a continuous performance task when offering performance-based rewards (both time and money).

Several studies have shown an attenuation of the vigilance decrement using specific incentives. For example, Sipowicz, Ware, and Baker (1962) had four groups complete a 3-hour vigilance task. One condition had a reward component, one had a feedback component, one had both feedback and a reward, and one condition served as a control. In each condition, participants had to monitor for brief offsets of a light, which occurred roughly once every 2.5 minutes. All experimental conditions detected signals at a higher rate than the control condition, and the reward + feedback condition performed significantly better than all other conditions. When looking at changes in performance across time, Sipowicz et al. (1962) found that the control group demonstrated a deterioration in signal detection, but none of the experimental conditions did. Thus, they effectively eliminated the vigilance decrement with feedback and reward. Esterman et al. (2016) had participants complete the gradCPT, a task that requires continuous discriminations between images of two categories, while the images gradually change from one to the next. In this task, discrimination (d') tends to decrease across time. Esterman et al. (2016) notified participants that earning a large reward (\$18) was contingent upon their performance on one particular trial, which could occur at any point during the task. If participants made an error on this trial, they would lose the entire reward. In these situations, participants did not show a vigilance decrement. Massar et al. (2019) also showed attenuation of the vigilance decrement in the psychomotor vigilance decrement with monetary rewards. Another study showed reductions in the vigilance decrement following intensive meditation training (MacLean et al., 2010). Finally, the vigilance decrement can be eliminated in the psychomotor vigilance task when intertrial intervals are short and predictable (Unsworth & Robison, 2020; 2018). These results suggest that, when sufficiently motivated, the vigilance decrement can be attenuated and even eliminated over both brief (~ 10 min) and long (~3 hr) intervals.

In the present study, we used a simple reaction time task to measure sustained attention. We chose this task (the psychomotor vigilance task; Dinges & Powell, 1985) because it is a relatively straightforward yet demanding task that requires continuous engagement. Because it is a simple reaction time task, it does not require any perceptual discriminations, nor is it susceptible to speed-accuracy tradeoffs or learning. We and others have used this task extensively to examine sustained attention in the laboratory (Brewer et al., 2017; Drummond et al., 2005; Massar et al., 2016, 2019; Unsworth & Robison, 2016, 2018, 2020; Unsworth et al., 2018). We compared distributions of reaction times as well as the vigilance decrement (i.e., slowing of reaction time) across conditions. Our reason for examining distributions of reaction times, rather than a summary score (e.g., mean or median), is that occasionally a meaningful experimental effect or individual difference can lie in a certain portion of the distribution, but not appear in the summary score (Balota & Yap, 2011). Additionally we collected several types of subjective data. Our hypotheses relied on the idea that goal-setting, feedback, and incentives would increase motivation and task engagement. To measure this, we inserted thought probes into the experiment to measure momentary attentional states. This technique has become a fairly common method of measuring the

degree to which various experimental manipulations affect mind-wandering and other off-task attentional states (Robison et al., 2019; Weinstein, 2018). If participants become more engaged with the task because of the experimental manipulations, then they should report fewer instances of task-unrelated thoughts (TUTs) and greater motivation. To measure motivation, we inserted a question at the end of each experimental block asking participants to report their current level of task-based motivation. We also asked participants to report their current level of alertness after each block, as well. These questions were included to measure subjective sensations of fatigue as participants proceeded through the task.

In Experiment 1, we focus on the effect of setting a specific goal on sustained attention. We compare conditions in which participants are given rather standard instructions with a vague goal to conditions that provide a specific, difficult goal and a specific, easy-to-attain goal, respectively. We hypothesized that providing a specific, difficult goal would improve sustained attention in the form of speeding reaction times and/or mitigating the vigilance decrement. In Experiment 2, we combined goal-setting and feedback, as feedback has been identified as a moderator of goals (Locke & Latham, 2002). We hypothesized that feedback would improve sustained attention, and further that feedback and goal-setting would interact, such that the effect of a goal would be magnified when people are given feedback as to how they are (or are not) meeting the goal. In Experiment 3, we offered incentives to reach a goal, as incentives have also been identified as moderators of goals (Locke & Latham, 2002). We hypothesized that providing incentives to reach a goal would improve sustained attention beyond simply providing the goal. Finally, in Experiment 4, we examined a broader array of goal difficulties to test the idea that optimal performance and motivation is achieved at moderately-difficult goals.

To summarize, the present set of experiments had several aims: 1) apply goal-setting theory to a sustained attention task to identify experimental manipulations that can mitigate or eliminate the vigilance decrement, 2) collect a multi-modal dataset that allows for a more well-rounded examination of the subjective experiences that accompany vigilance tasks (i.e., mind-wandering, motivation, fatigue), and 3) develop testable hypotheses for future work regarding theoretical accounts of vigilance.

Experiment 1

The goal of Experiment 1 was to determine whether setting specific goals would improve sustained attention. Based on goal-setting theory, we hypothesized that setting a specific, difficult goal would improve task performance, increase motivation, and decrease task-unrelated thoughts. Thus we provided participants in one condition with a rather difficult goal to try to attain, we provided participants in a second condition a rather easy goal to attain, and we included one condition with standard, vague instructions. We set the easy and difficult goal standards based on a distribution of performance from this task in prior work (Unsworth & Robison, 2016).

Method

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Participants and Procedure

A sample of 105 participants from the human subjects pool at the University of Oregon completed the study in exchange for partial course credit. One participant was removed from the analyses for having more than 10 excluded RTs in their data (see Results), leaving a final sample of 104 participants for the analyses (74 women, 30 men, $M_{age} = 19.16$, $SD_{age} = 1.93$). The session lasted 1 hr, the first 30 minutes of which included the sustained attention task. During the remaining 30 minutes of the session, participants completed a visual working memory task, the results of which have been reported elsewhere (Robison & Unsworth, 2019). Participants completed an informed consent and brief demographic questionnaire before starting the session. The experimental protocol was approved by the Institutional Review Board of the University of Oregon. At the outset of the experiments, we established a target sample size of 35 participants per condition.¹

Task

Participants completed a 30-minute version of the psychomotor vigilance task. Each trial began with a 2-s fixation screen (++++). Then, a row of blue zeroes appeared at the center of the screen (00.000). After a random time interval (2 - 10 s in 500-ms intervals), the numbers began counting up like a stopwatch. The participants' task was to stop the numbers from counting as quickly as possible by pressing the spacebar. After the participant pressed the spacebar, the numbers stopped counting, and the RT (e.g, 00.356) remained on-screen in red font for 1 s. The next trial began after a 1-s blank screen. The task was divided into 5 blocks of 28 trials (140 total trials). After the 28 trials within a block, the task stopped and asked participants to make a motivation and alertness rating. Previous research using this task of similar duration has shown that RTs tend to increase across time, as do reports of task-unrelated thoughts (Unsworth & Robison, 2016).

Motivation and alertness ratings

At the end of each block, participants were asked to rate their current level of motivation and alertness on a 7-point scale. The motivation question asked, "Please rate how motivated you feel right now to perform well on the task." The anchors were 1 = "not at all" and 7 = "very motivated". The alertness measure said, "Please rate how alert you feel right now." The anchors were 1 = "not at all" and 7 = "very alert."

¹This sample size was set a priori during a dissertation proposal by author M. K. R.

Thought probes

Periodic thought-probe screens asked the participants to indicate their momentary attentional state. After four randomly-sampled trials in each block, a screen appeared saying, “Press the key that best describes what you were thinking about *just prior* to this screen appearing.” Participants were given six response options: 1) I was totally focused on the current task, 2) I was thinking about my performance on the task, 3) I was distracted by sights/sounds in my environment, 4) I was intentionally thinking about things unrelated to the task, 5) I was unintentionally thinking about things unrelated to the task, and 6) My mind was blank. We recorded response 1 as on-task, response 2 as task-related interference, and responses 3 - 6 as task-unrelated thought (TUT; Stawarczyk, Majerus, Maj, Van der Linden, & D’Argembeau, 2011).

Goal-setting instructions

The experiment included three between-subjects conditions: no-goal, hard goal, and easy goal. In the no-goal condition, participants were instructed, “Previous research has demonstrated that good performance on this task is a fast reaction time. So your goal should be to keep your average reaction time as low as possible.” In the hard-goal condition, participants were instructed, “Previous research has demonstrated that good performance on this task is a reaction time below 300 milliseconds. So your goal should be to keep your average reaction time below 300 milliseconds.” In the easy goal condition, participants were instructed, “Previous research has demonstrated that good performance on this task is a reaction time below 800 milliseconds. So your goal should be to keep your average reaction time below 800 milliseconds.”

Post-experiment questionnaire

At the end of the task, participants were given a sheet of paper with the question, “What was your goal on this task?” Participants wrote their responses under the question in an open-ended fashion. Participants were also asked, “Before you came to the lab for this study, had you heard anything (from classmates, friends, etc.) about this specific study?” Participants checked “Yes” or “No.” If they responded Yes, they were asked to report what they knew about the study coming into the lab.

Eye-tracking

In Experiments 1 - 3, participants completed the task in a dark room in front of an eye-tracker with their head position fixed in a chinrest. The eye-tracker was used to collect pupillometry data while participants completed the task. The interested reader can see these data in Robison (2018).

Data analysis

We used R for all our analyses. All analysis code can be found in an R Markdown document on the Open Science Framework at the following url: osf.io/n7qa2/. We trimmed RTs outside 200 and 3000 ms so the results would not be skewed by anticipatory responses or abnormally long RTs. We then removed any participants who had 10 or more of these RTs. We used the *tidyverse* package (Wickham, 2017) to aggregate and plot data, the *gridExtra* package (Auguie, 2017) to combine plots, the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) to fit linear mixed effects models, the *EMAtools* package (Kleiman, 2017) to estimate effect sizes for the models, the *BayesFactor* (Morey & Rouder, 2018) package to compare models and estimate Bayes Factors, and the *papaja* package (Aust & Barth, 2018) to assemble the manuscript in R Markdown.

To assess the balance of evidence in favor of (or against) a particular effect or interaction, as opposed to traditional null hypothesis statistical testing, we compared nested models using Bayes Factors (BFs). The BF can be interpreted as the ratio of evidence in favor of a particular model over another model. Here we always compared nested models, computing the ratio of evidence in favor of including a particular effect or interaction compared to a model without this effect (listed in the tables as BF_{10}). We interpreted BFs > 3 as strong evidence for an effect, BFs $< .33$ as strong evidence against an effect, and BFs between $.33$ and 3 as indeterminate. BFs for main effects are compared against an intercept-only model (only a random intercept for each participant). BFs for interactions are compared to a model with main effects but no interaction. Effect sizes are the estimate from the fully-parameterized model. By reporting BFs in combination with standardized regression coefficients and effect size (Cohen's d), our hope was to estimate the magnitude of a given effect and the weight of evidence for or against our hypothesis given the data.

Results & Discussion

We ran two sets of analyses to test our two hypotheses that 1) setting a specific goal would improve sustained attention compared to a non-specific goal, and 2) setting a difficult goal would improve sustained attention compared to an easy goal. To do so we created two separate contrasts in the data: one comparing the two goal conditions to the no-goal condition and one comparing the easy goal condition to the hard goal condition. Then, for each dependent variable, we specified two sets of linear mixed effect models - one with goal specificity as the independent variable, and one with goal difficulty as the independent variable (*Note*: participants in the no-goal condition were not included in the latter analysis). To estimate the magnitude of effects, we used Cohen's d . To estimate the weight of evidence in favor of (or against) an effect or interaction, we used Bayes Factors (BFs). In the tables, the BFs listed are the ratio of evidence in favor of including a given main effect or interaction compared to a model that excludes it, and the d values are the effect size in the full model including all main effects and interactions.

We analyzed RTs in two ways. In the first, we examined distributions of RTs by rank-ordering each individual's RTs and binning them into quintiles from fastest to slowest. This allowed us to examine the experimental effects on the full distribution of RTs, rather

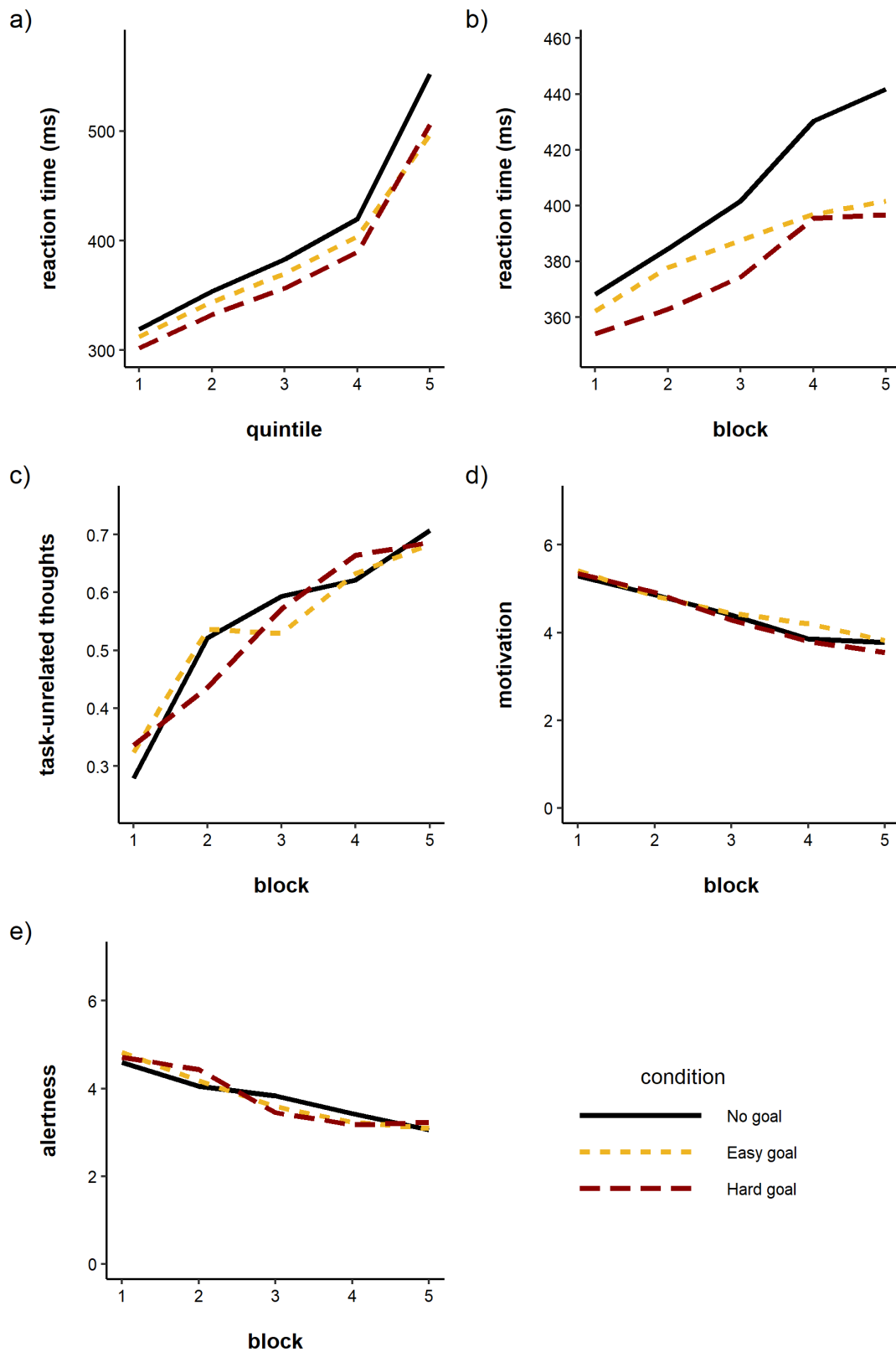


Figure 1. a) Reaction times, b) task-unrelated thoughts, c) motivation ratings, and d) alertness ratings by block and condition in Experiment 1.

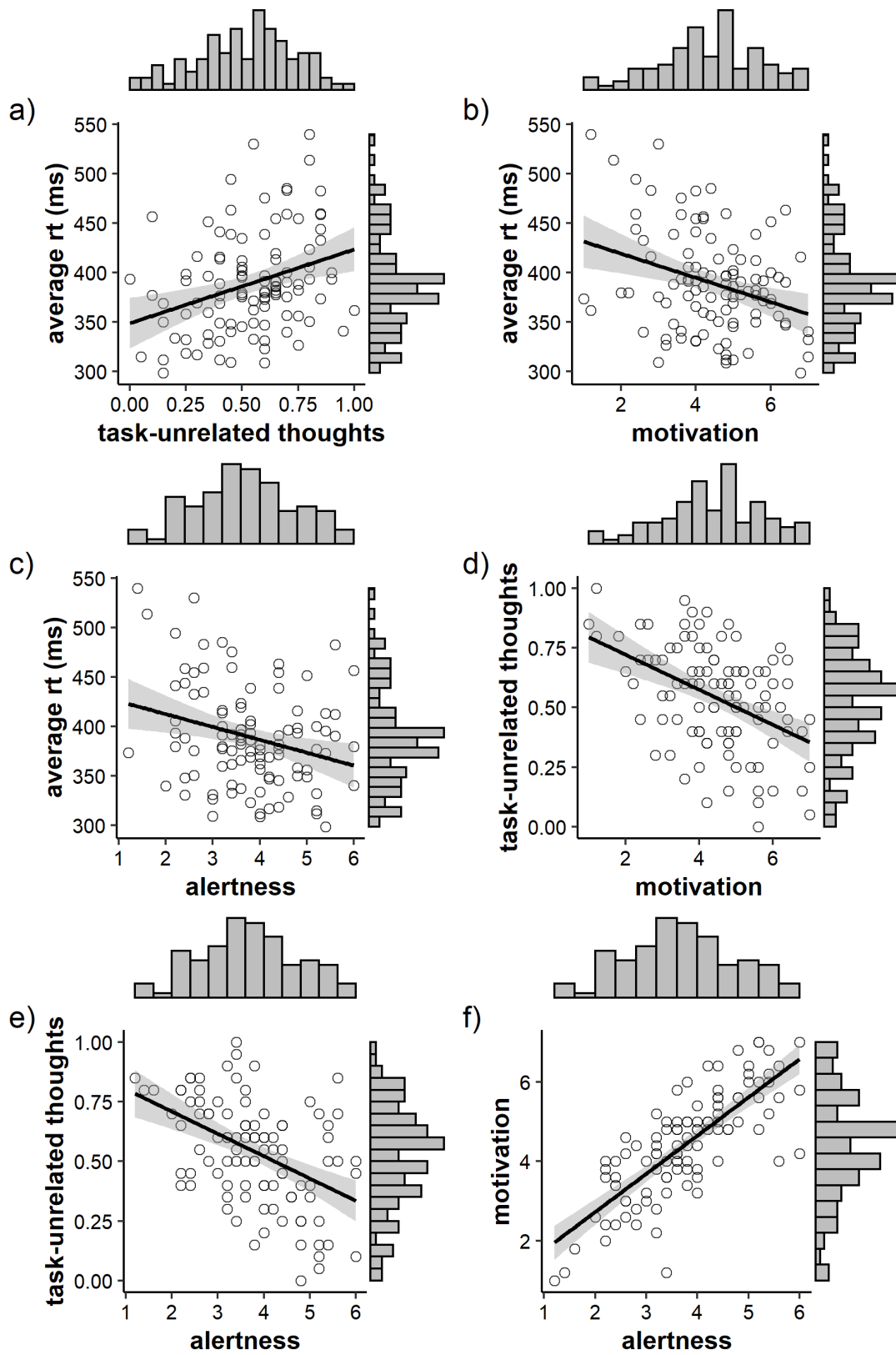


Figure 2. Scatterplots of correlations among dependent variables in Experiment 1.

Table 1
Coefficients of full models with goal-specificity as a factor in Experiment 1

DV	Effect	B	SE	<i>t</i>	<i>d</i>	BF ₁₀
Reaction times (quintile)	Bin	0.40	0.01	64.28	1.07	> 100,000
	Goal	0.01	0.08	0.10	0.02	1.26
	Bin x Goal	-0.06	0.01	-8.36	-0.14	> 100,000
Reaction times (block)	Block	0.14	0.01	18.15	0.30	> 100,000
	Goal	-0.18	0.08	-2.33	-0.46	1.27
	Block x Goal	-0.06	0.01	-6.45	-0.11	> 100,000
Task-unrelated thoughts	Block	0.10	0.01	8.05	0.36	> 100,000
	Goal	0.00	0.05	-0.10	-0.02	0.06
	Block x Goal	-0.01	0.01	-0.51	-0.02	0.09
Motivation	Block	-0.24	0.03	-7.90	-0.78	> 100,000
	Goal	0.01	0.16	0.09	0.02	0.23
	Block x Goal	-0.01	0.04	-0.37	-0.04	0.11
Alertness	Block	-0.26	0.03	-8.18	-0.80	> 100,000
	Goal	0.00	0.16	-0.01	0.00	0.20
	Block x Goal	-0.04	0.04	-1.08	-0.11	0.16

Note. DV = dependent variable, B = standardized regression coefficient, SE = standard error of regression coefficient, BF₁₀ = Ratio of evidence (Bayes Factor) in favor of effect.

than a summary score (e.g., mean RT), as an experimental manipulation may affect the full distribution of RTs, or it may have a selective effect on different parts of the distributions (e.g. the slowest RTs; Balota & Yap, 2011). Then, we examined changes in RT across the duration of the task. This allowed us to examine how the experimental manipulations affected the vigilance decrement. For the remaining dependent variables, we used this same technique. To do so, we specified linear mixed effect models with fixed effects of block and goal specificity/difficulty and a random effect of participant. The intercept was allowed to vary across participants.² For the distributional analysis of RTs, we specified a linear mixed model with goal specificity/difficulty and quintile as fixed effects and participant as a random effect. Model summaries of the models with goal *specificity* as a factor are listed in Table 1, and model summaries with goal *difficulty* as a factor are listed in Table 2. The data are plotted in Figure 1.

When comparing distributions of RTs across the goal and no-goal conditions, the model revealed a large effect of bin and a modest bin x goal interaction. The model comparisons using BFs revealed indiscriminate evidence regarding the main effect of goal specificity (Table 1). The model comparisons also indicated that there was strong evidence for a goal x bin interaction. What this indicates is that the largest effect of goal specificity on RTs manifested in the slowest reaction times (see Figure 1). Often, these slow reaction times

²We attempted to estimate mixed models with a random slope (i.e., effect of block) as well. But some of these models did not converge upon a solution. Therefore, we only allowed the intercept to vary.

Table 2
Coefficients of full models with goal-difficulty as a factor in Experiment 1

DV	Effect	B	SE	<i>t</i>	<i>d</i>	BF ₁₀
Reaction times (quintile)	Bin	0.32	0.01	62.80	1.28	> 100,000
	Difficulty	-0.06	0.08	-0.78	-0.19	0.18
	Bin x Difficulty	0.02	0.01	3.23	0.07	0.13
Reaction times (block)	Block	0.07	0.01	10.49	0.21	> 100,000
	Difficulty	-0.06	0.08	-0.79	-0.19	0.26
	Block x Difficulty	0.01	0.01	1.52	0.03	0.13
Task-unrelated thoughts	Block	0.08	0.01	6.74	0.37	> 100,000
	Difficulty	0.00	0.05	-0.09	-0.02	0.13
	Block x Difficulty	0.01	0.02	0.73	0.04	0.12
Motivation	Block	-0.23	0.03	-6.97	-0.84	> 100,000
	Difficulty	-0.10	0.20	-0.48	-0.12	0.32
	Block x Difficulty	-0.05	0.05	-1.20	-0.15	0.24
Alertness	Block	-0.31	0.03	-8.88	-1.07	> 100,000
	Difficulty	0.01	0.17	0.07	0.02	0.22
	Block x Difficulty	0.01	0.05	0.26	0.03	0.12

Note. DV = dependent variable, B = standardized regression coefficient, SE = standard error of regression coefficient, BF₁₀ = Ratio of evidence (Bayes Factor) in favor of effect.

Table 3
Correlations among dependent variables in Experiment 1

	1	2	3	4
1. Reaction time	–			
2. Task-unrelated thoughts	.32	–		
3. Motivation	-.32	-.45	–	
4. Alertness	-.27	-.46	.78	–

are interpreted as attentional lapses, moments when the individual’s attention has strayed from the task leading to a very slow reaction. It appears setting a specific goal reduced the severity of such instances.

Comparing dependent variables across goal and no-goal conditions with block as a factor revealed a large effect of time on all the dependent variables. Across blocks, RTs and TUTs increased and motivation and alertness decreased. These findings replicate the typical empirical pattern that sustained attention deteriorates across time (Esterman & Rothlein, 2019; Hopstaken et al., 2015a), and more specific patterns for this task (Unsworth & Robison, 2016). With regard to the vigilance decrement, there was strong evidence for a block x goal interaction on RTs. Thus, setting a specific goal attenuated the vigilance decrement (see Figure 1). Despite these effects on RTs, the goal conditions did not produce significantly fewer TUTs, higher motivation ratings, or higher alertness ratings. In fact, the BFs indicated fairly strong evidence *against* the presence of such effects. So although there was evidence that the goal-setting manipulations did improve sustained attention in form of reduced severity of long RTs and an attenuated vigilance decrement, the specific goals did not affect the subjective levels of task engagement reported by participants.

Next we ran the same set of analyses but with goal *difficulty* as a factor rather than goal *specificity*. The effect of goal difficulty on RT was rather modest in magnitude, and the model comparison favored the absence of this effect. For all other measures, there was evidence against the hypothesis that a more difficult goal would lead to more task engagement than an easier goal (see Table 2).

Taken together, the results confirmed one set of hypotheses - that setting a specific goal would improve sustained attention in the form of a reduced vigilance decrement and reduced skew in their longest RTs. However, there was evidence against our hypothesis that setting a difficult goal would do the same. Further, we did evidence against our hypothesis a specific goal would improve subjective task engagement. This begged the question - are the task engagement measures even valid? To answer this question, we examined correlations among the measures to assess whether the task engagement measures were meaningfully related to task performance. These correlations are listed in Table 3 and depicted visually in Figure 2. Indeed, the subjective measures correlated with task performance meaningfully (all $|r|s > .27$). Participants who reported fewer TUTs, higher motivation, and higher alertness tended to have better task performance (i.e., faster RTs). Further, participants who reported higher motivation tended to report fewer TUTs. Therefore, there were meaningful relations between task performance and task engagement, indicating that these measures were capturing important interindividual variance.³

There are several reasons why the goal-setting manipulation did not affect the subjective measures of task engagement. First, it could be the case that the goals were not paired with performance feedback. The only feedback participants received was their RT on each trial. Therefore the lack of feedback could have reduced the effectiveness of the goal-setting instructions. We test this hypothesis in Experiment 2. Second, the goal was not

³Because participants are in different experimental conditions, some of the interindividual variance will be due to their condition-specific instructions. However, partialing condition out of the correlations did not change them, so we report raw correlations here.

paired with any significant reward. Therefore, there was nothing to encourage participants to commit to the stated goal. We test whether pairing the goal with a specific reward has a moderating effect in Experiment 3. Third, it could be the case that the hard goal we specified was indeed *too* difficult. Some theories argue that moderately difficulty goals lead to optimal behavior (e.g, Atkinson, 1957). This is a possibility we examine in Experiment 4.

Experiment 2

Although Experiment 1 did find an effect of goal-setting on the vigilance decrement, there was perhaps one crucial element missing – feedback. The instructions provided a goal, and participants saw their RT after each trial. So they knew on each trial whether their RT fell above or below the goal. But it may have been difficult for participants to keep a running average of their RTs, and the stated goal was to keep their average RT below a specific threshold (or as low as possible). So in Experiment 2, we provided block-by-block feedback to inform participants about their performance. And, when given a goal, we provided feedback as to whether or not participants were currently meeting the goal. In a 2 x 2 design, we crossed goal-setting and feedback in a fully between-subjects design. In the no-goal + no-feedback condition, participants were given neither a specific goal nor performance feedback. In the goal + no-feedback condition, participants were given the same instructions as the hard-goal condition from Experiment 1, but they were given neither trial-by-trial nor block-by-block feedback about their performance. In the no-goal + feedback condition, participants were given the same instructions as the no-goal condition from Experiment 1, but they were given both trial-by-trial and block-by-block feedback. Finally, in the goal + feedback condition, participants were given the hard-goal instructions from Experiment 1 and trial-by-trial and block-by-block feedback.

Locke and Latham (2002) propose that feedback can be a moderating factor with regard to goals. That is, providing feedback as to how people are or are not meeting a goal standard can make goals more effective. Several studies have shown improvement in simple reaction time experiments by providing participants with feedback (also called knowledge-of-results; Church & Camp, 1965; McCormack, 1959; McCormack, Binding, & Chylinski, 1962; McCormack, Binding, & McElheran, 1963; Sipowicz et al., 1962; Warm, Riechmann, Grasha, & Seibel, 1973). Some have argued that feedback is only effective when it is paired with a specific goal (Locke & Latham, 2002; Locke, Shaw, Saari, & Latham, 1981). But others have gone further to say that goals *must* be paired with feedback in order to be effective (Erez, 1977; Payne & Hauty, 1955; Strang, Lawrence, & Fowler, 1978). However, feedback might also be beneficial even in the absence of a specific goal. A few studies have specifically tested whether feedback affects the vigilance decrement, as opposed to overall levels of performance, and they have differing findings. Sipowicz et al. (1962) found an elimination of the vigilance decrement when participants were given feedback or a combination of feedback and a reward. But Warm et al. (1973) found only a main effect of feedback, not an attenuation or elimination of the vigilance decrement. Theoretically, feedback can provide scaffolding for the naturally-existing self-regulation mechanisms people possess. In general, people monitor their behavior and adjust in response to feedback they

receive from their environments. Explicit feedback can provide additional scaffolding for people to regulate their behavior.

We hypothesized that providing feedback would have a beneficial effect on task performance in the form of reduced reaction times and/or a shallower vigilance decrement, fewer task-unrelated thoughts, and greater motivation. We also hypothesized that feedback and goal-setting would interact. That is, we expected feedback to have a larger effect when it was paired with a specific, difficult goal than when it was provided without a goal. By the same token, we expected the goal-setting instructions to have more of an impact when the goal was paired with feedback.

Method

Participants and Procedure

A sample of participants from the human subjects pool at the University of Oregon completed the study in exchange for partial course credit. The session lasted 1 hr, the first 30 minutes of which included the sustained attention task. During the remaining 30 minutes of the session, the participants completed a visual working memory task, the results of which have been reported elsewhere (Robison & Unsworth, 2019). We set a target sample size of 35 participants per condition. Three participants did not understand the instructions and one participant fell asleep during the session, leaving a final sample of 139 participants in the analyses (89 women, 50 men, $M_{age} = 19.60$, $SD_{age} = 2.33$).

Psychomotor vigilance task

The task was identical to that used in Experiment 1 with the exception of the addition and removal of feedback for the feedback and no-feedback conditions, respectively.

Motivation and alertness ratings

See Experiment 1.

Thought probes

See Experiment 1.

Goal-setting and feedback

The 2 x 2 between-subjects design resulted in four conditions with goal-setting and feedback crossed across conditions. In the goal conditions, participants were given the same instructions as the hard-goal condition from Experiment 1: “Previous research has demonstrated that good performance on this task is a reaction time below 300 milliseconds. Therefore you should try to keep your average reaction time below 300 milliseconds.” In the

no-goal conditions, participants were given the same instructions as the no-goal condition from Experiment 1: “Previous research has demonstrated that good performance on this task is a fast reaction time. Therefore you should try to keep your average reaction time as low as possible.”

In the no-feedback conditions, we removed the trial-by-trial feedback and did not provide block-by-block feedback. Instead of seeing their reaction time for 1 s following each trial, the screen showed red Xs (XX.XXX) for 1 s. At the end of each block, a screen appeared that said, “That is the end of this block of trials. The next block is about to begin.” This remained on the screen for 8 s, then the next block of trials began. In the feedback conditions, we retained the trial-by-trial feedback and we added block-by-block feedback as well. At the end of each block, a screen appeared saying, “That is the end of this block of trials. Your average reaction time for the previous block was XXX ms. Your average reaction time for the experiment is XXX ms.” In the goal + feedback condition, an additional sentence appeared below the feedback saying, “You are currently meeting the goal, keep it up!” if their average reaction time for the experiment was below 300 ms. If their average reaction time was above 300 ms, the screen said, “You are not currently meeting the goal. Try to respond faster!” The feedback screen remained for 8 s before the next block of trials began. The feedback screens appeared after the motivation and alertness rating screens.

Post-experiment questionnaire

See Experiment 1.

Results & Discussion

We took the same analytic approach as we did in Experiment 1. The models on each dependent variable are summarized in Tables 4 and 5, and the data are plotted in Figure 3. There was strong evidence for bin x goal, bin x feedback, and bin x goal x feedback interactions. As can be seen in Figure 3a, the combination of a goal with feedback particularly reduced the longest RTs in the distribution. There was also strong evidence for a block x goal interaction, which indicates that giving a specific goal attenuated the vigilance decrement, replicating Experiment 1. Additionally, there was strong evidence in favor of a block x goal x feedback interaction, indicating that participants in the goal + feedback condition demonstrated the shallowest vigilance decrements (Figure 3b). The evidence for a main effect of feedback and a block x feedback interaction on RTs was indiscriminate.

Similar to Experiment 1, there were rather large effects of time on all the other dependent variables. There was strong evidence for a main effect of feedback on TUTs. Participants who received feedback reported fewer TUTs. There was evidence against a block x goal, block x feedback, and block x goal x feedback interactions, and the data were indiscriminate regarding a main effect of goals and the goal x feedback interaction. Regarding the motivation and alertness ratings, there was evidence against a main effect of goal, the block x feedback interaction, and the block x goal x feedback interaction, and the data were indiscriminate regarding a main effect of feedback, a block x goal interaction, and

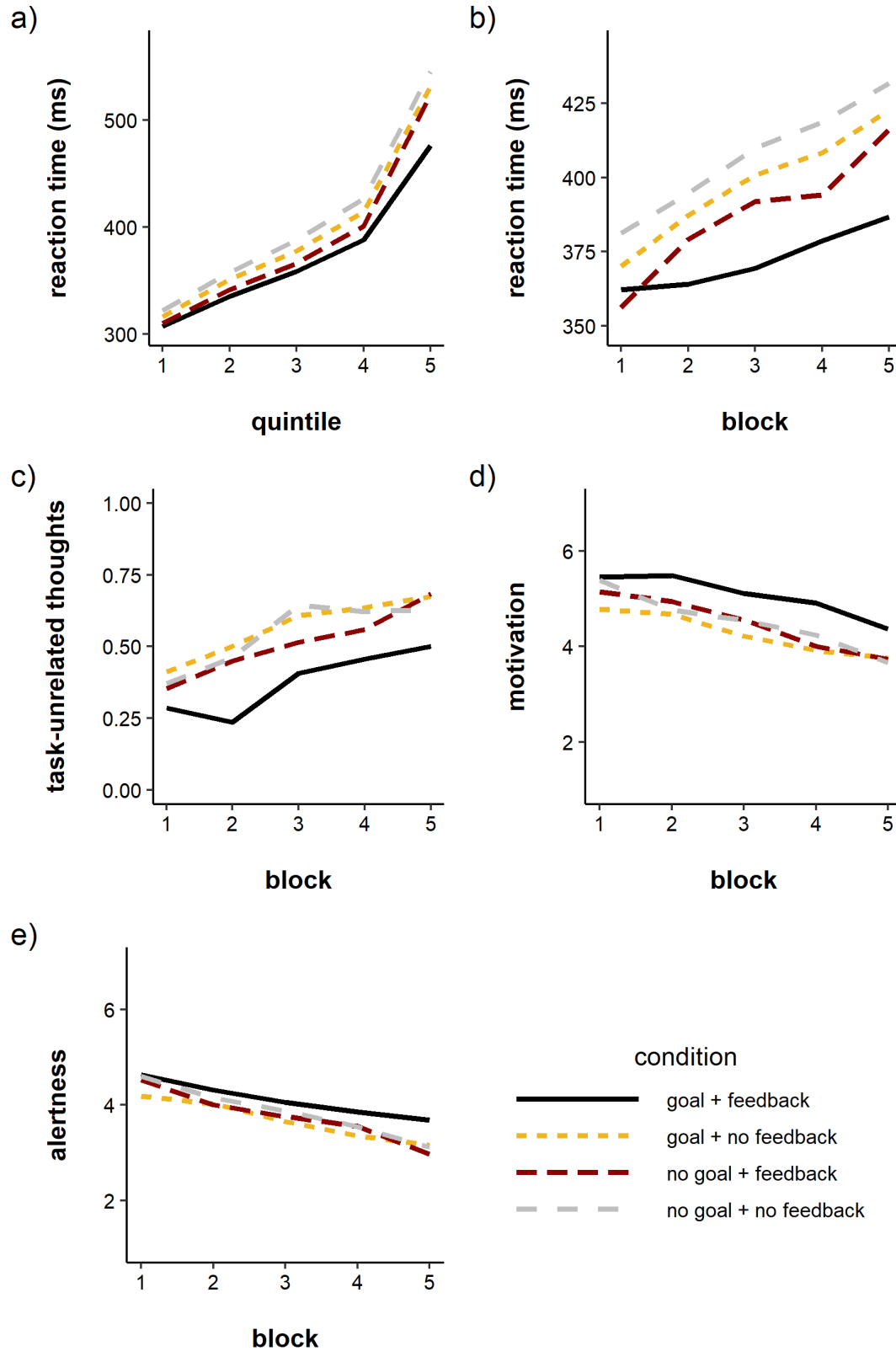


Figure 3. a) Reaction times, b) task-unrelated thoughts, c) motivation ratings, and d) alertness ratings by block and condition in Experiment 2.

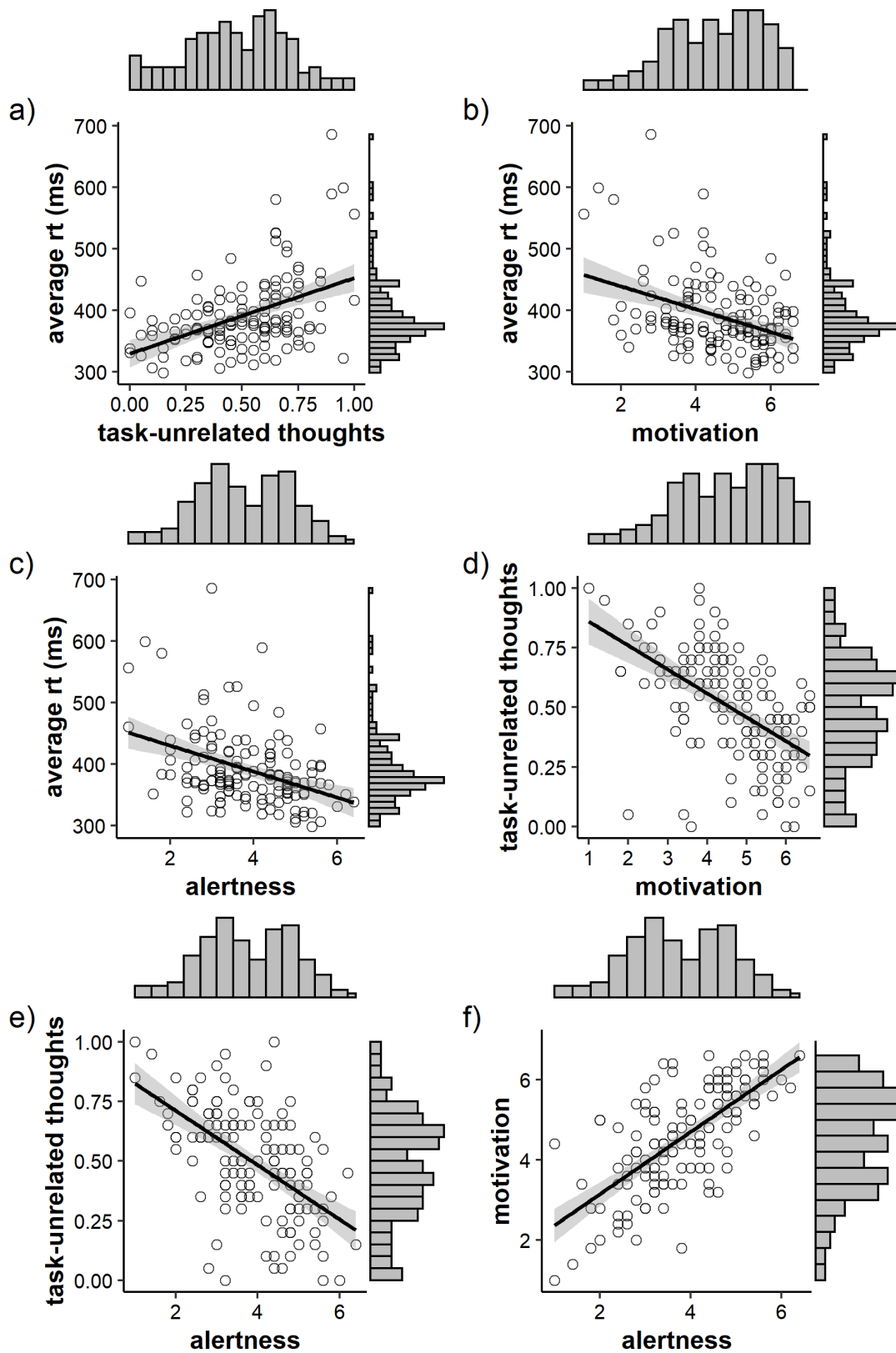


Figure 4. Scatterplots of correlations among dependent variables in Experiment 2.

Table 4
Summaries of linear mixed effects models in Experiment 2

DV	Effect	B	SE	t	d	BF_{10}
Reaction times (quintile)	Bin	0.42	0.01	59.13	0.85	> 100,000
	Goal	-0.08	0.12	-0.63	-0.11	0.20
	Feedback	-0.16	0.12	-1.32	-0.23	1.00
	Bin x Goal	-0.02	0.01	-1.56	-0.02	> 100,000
	Bin x Feedback	-0.03	0.01	-2.63	-0.04	> 100,000
	Goal x Feedback	-0.05	0.17	-0.28	-0.05	0.18
	Bin x Goal x Feedback	-0.06	0.01	-4.47	-0.06	354.22
Reaction times (block)	Block	0.10	0.01	11.38	0.16	> 100,000
	Goal	-0.08	0.12	-0.65	-0.11	0.20
	Feedback	-0.16	0.12	-1.33	-0.23	1.01
	Block x Goal	0.00	0.01	0.12	0.00	3.17
	Block x Feedback	0.01	0.01	0.64	0.01	0.57
	Goal x Feedback	-0.05	0.17	-0.28	-0.05	0.13
	Block x Goal x Feedback	-0.06	0.02	-3.39	-0.05	10.82
Task-unrelated thoughts	Block	0.20	0.04	5.02	0.43	> 100,000
	Goal	0.06	0.16	0.40	0.07	0.39
	Feedback	-0.10	0.16	-0.63	-0.11	9.37
	Block x Goal	0.00	0.06	-0.07	-0.01	0.13
	Block x Feedback	0.03	0.06	0.52	0.04	0.11
	Goal x Feedback	-0.47	0.23	-2.09	-0.36	1.79
	Block x Goal x Feedback	-0.03	0.08	-0.42	-0.04	0.17

a goal x feedback interaction on motivation. Similarly, the analysis on alertness revealed evidence against a main effect of goal, a main effect of feedback, a block x feedback interaction, and a block x goal x feedback interaction. The data were indiscriminate regarding the block x goal interaction and the goal x feedback interaction. Thus the only strong effect we observed in the subjective reporting data was on TUT, with subjects in the feedback conditions reporting fewer TUTs than the no-feedback conditions. Finally, we examined interindividual correlations among RT, TUT, motivation, and alertness. The measures all meaningfully correlated with one another (all $|r|s > .38$), similar to Experiment 1 (see Table 6 and Figure 4). Thus there were relations among the measures of task performance and task engagement.

The results created two additional follow-up questions. First, was the effect of feedback due to the addition of block-by-block feedback, the removal of trial-by-trial feedback, or both? Second, what type(s) of TUT did goals and feedback reduce? We can answer the first question by comparing the goal + feedback condition in Experiment 2 to the hard goal condition in Experiment 1. These two conditions differed in only one respect: the goal + feedback condition in Experiment 2 received block-by-block feedback, and the hard goal condition in Experiment 1 did not. Both received trial-by-trial feedback. If these two conditions differ, then we can conclude that the *addition* of block-by-block feedback helped

Table 5
Summaries of linear mixed effects models in Experiment 2 (continued)

DV	Effect	B	SE	<i>t</i>	<i>d</i>	BF ₁₀
Motivation	Block	-0.26	0.03	-8.53	-0.73	> 100,000
	Goal	-0.16	0.19	-0.84	-0.14	0.29
	Feedback	-0.03	0.20	-0.15	-0.03	0.94
	Block x Goal	0.07	0.04	1.81	0.15	1.57
	Block x Feedback	0.01	0.04	0.31	0.03	0.09
	Goal x Feedback	0.54	0.27	1.99	0.34	1.80
	Block x Goal x Feedback	-0.01	0.06	-0.15	-0.01	0.13
Alertness	Block	-0.25	0.03	-8.18	-0.70	> 100,000
	Goal	-0.13	0.19	-0.68	-0.12	0.23
	Feedback	-0.07	0.20	-0.35	-0.06	0.32
	Block x Goal	0.06	0.04	1.41	0.12	1.47
	Block x Feedback	0.00	0.04	0.03	0.00	0.10
	Goal x Feedback	0.38	0.27	1.37	0.24	0.86
	Block x Goal x Feedback	0.03	0.06	0.43	0.04	0.16

Note. DV = dependent variable, B = standardized regression coefficient, SE = standard error of regression coefficient, BF₁₀ = Ratio of evidence (Bayes Factor) in favor of effect.

Table 6
Correlations among dependent variables in Experiment 2

	1	2	3	4
1. Reaction time	–			
2. Task-unrelated thoughts	.46	–		
3. Motivation	-.38	-.56	–	
4. Alertness	-.39	-.56	.69	–

participants in the goal + feedback condition. Comparing these two conditions revealed strong evidence for a reduction in the vigilance decrement on RTs ($B = -0.02$, $BF_{10} = 178.64$), a smaller effect of time on motivation ($B = 0.06$, $BF_{10} = 3.19$), and a smaller effect of time on alertness ($B = 0.07$, $BF_{10} = 6.47$). The evidence was indiscriminate regarding an effect on TUT ($B = -0.11$, $BF_{10} = 0.43$). Thus, pairing block-by-block feedback with a goal clearly had a facilitative effect over and above combining that same goal with only trial-by-trial feedback.

We can also directly compare the hard goal condition from Experiment 1 to the goal + no-feedback condition from Experiment 2. These conditions differed only in the presence/absence of trial-by-trial feedback. If these conditions differ, we can also conclude the *removal* of trial-by-trial feedback was actually harmful, thus leading to an effect of feedback in Experiment 2. There was strong evidence against an effect of trial feedback on

the vigilance decrement ($B = 0.00$, $BF_{10} = 0.05$). The trial feedback did reduce TUTs (main effect: $B = -0.16$, $BF_{10} = 5.25$; block x feedback interaction: $B = 0.38$), but it did not affect motivation (main effect: $B = 0.19$, $BF_{10} = 0.28$), or alertness ($B = 0.20$, $BF_{10} = 0.27$). Thus, it appears the effects of feedback observed in Experiment 2 were primarily driven by the presence/absence of block-by-block feedback rather than trial-by-trial feedback.

We can answer the second question (i.e., which types of TUTs were affected?) by separately examining each TUT response (external distraction, intentional mind-wandering, unintentional mind-wandering, and mind-blanking). The goal-setting and feedback manipulations did not have an effect on external distraction reports, which were rare (6.99% of all reports; all BF_{10} s < .14). Unintentional mind-wandering was quite commonly reported (28.29% of reports), but the manipulations did not affect it (all BF_{10} s < .24). Intentional mind-wandering was rarely reported (28.29% of reports), but feedback significantly reduced it ($B = -0.03$, $BF_{10} = 10.50$). Finally, mind-blanking (9.23% of reports) was unaffected by the manipulations (all BF_{10} s < .40). Thus, the effect we observed on TUTs was mainly observed because the feedback manipulations reduced intentional mind-wandering, rather than other types of TUTs. Interestingly, participants in the feedback conditions, particularly the goal + feedback condition, also reported more task-related interference, as evidenced by a main effect of feedback ($B = 0.16$, $BF_{10} = 3.44$) and goal x feedback interaction on these reports ($B = 0.03$, $BF_{10} = 56.13$). Thus participants who received feedback also tended to report thinking about their performance on the task more often, particularly when that feedback was tied to a specific goal.

Collectively, the data showed rather convincing evidence that combining a specific goal with can mitigate the vigilance decrement, in support of our hypothesis. Feedback also produced fewer TUTs, mostly due to a reduction in intentional mind-wandering. Importantly, feedback was a powerful modulator of TUTs regardless of whether or not it was paired with a specific goal, which is consistent with prior work (Church & Camp, 1965; McCormack, 1959; McCormack et al., 1962, 1963; Sipowicz et al., 1962; Warm et al., 1973).

Experiment 3

Experiment 3 examined whether providing a particular incentive to reach a goal would magnify the effect of a goal. Thus, we included the difficult goal condition with feedback, similar to the goal + feedback condition in Experiment 2, and two incentivized conditions. Theoretically, incentives work as a moderator of the effect of goals. If individuals perceive the goal as more worthy of their effort, they will be more committed to the goal. So in addition to collecting all the same dependent variables as in Experiments 1 and 2, we added one additional dependent variable: goal commitment. If incentives have an effect on sustained attention above and beyond the effect of a specific, difficult goal, they should do so via goal commitment (Locke & Latham, 2002).

Method

Participants and procedure

A sample of participants from the human subjects pool at the University of Oregon completed the study in exchange for partial course credit. Participants were randomly assigned to one of three conditions: no incentive, cash incentive, and time incentive. Regardless of condition, the experiment lasted about 30 min. We set a target sample of 35 participants per condition. Six participants indicated on the post-experimental questionnaire that they had heard about the study, and one participant felt sick in the middle of the experiment and had to take a break. These participants were replaced. The final sample included 111 participants (68 women, 42 men, 1 non-binary gender, $M_{age} = 19.39$, $SD_{age} = 2.08$).

Incentive manipulation

All participants were recruited to the study under the impression that the study would last 1 hr, and that they would receive 1 course credit for their participation. In all three conditions, participants were told, “Previous research suggests that good performance on this task is a reaction time below 300 milliseconds. Therefore, your goal on this task will be to keep your average reaction time below 300 milliseconds. After each block of trials, you will see a screen with your average reaction time for that block, your average reaction time overall, and how your performance compares to the goal.” The no-incentive condition was identical to the goal + feedback condition in Experiment 2. In the time-incentive condition, participants were told, “This task will take 1 hr. At the midway point of the task, if you have met the goal (i.e., kept our average reaction time below 300 ms), you will be allowed to stop the experiment early and leave the session. You will still receive 1 credit for participating today. If you have not met the goal, you will continue the task for another 30 minutes.” In the cash-incentive condition, participants were told, “At the end of the task, if you have met the goal (i.e., kept your average reaction time below 300 ms), you will receive a \$10 cash bonus in addition to 1 credit for your participation. If you have not met the goal, you will only receive course credit.” All participants who kept their average reaction time below 300 ms were given the \$10 bonus, regardless of condition. All participants were released after 30 minutes, regardless of performance, and given course credit for 1 hour of participation.

Psychomotor vigilance task

See Experiment 1.

Thought probes

See Experiment 1.

Motivation and alertness ratings

See Experiment 1.

Table 7

Summaries of models comparing time and no-incentive conditions in Experiment 3

DV	Effect	B	SE	<i>t</i>	<i>d</i>	BF ₁₀
Reaction times (quintile)	Bin	0.38	0.01	52.45	1.06	> 100,000
	Condition (time)	-0.04	0.12	-0.36	-0.08	0.14
	Bin x Condition (time)	0.01	0.01	1.43	0.03	0.04
Reaction times (block)	Block	0.07	0.01	8.24	0.17	> 100,000
	Condition (time)	0.01	0.11	0.05	0.01	0.14
	Block x Condition (time)	-0.01	0.01	-0.91	-0.02	0.05
Task-unrelated thoughts	Block	0.05	0.01	4.64	0.25	> 100,000
	Condition (time)	-0.07	0.05	-1.23	-0.29	0.21
	Block x Condition (time)	0.01	0.02	0.77	0.04	0.10
Motivation	Block	-0.19	0.03	-6.20	-0.73	> 100,000
	Condition (time)	0.38	0.18	2.09	0.50	1.63
	Block x Condition (time)	0.04	0.04	0.81	0.10	0.17
Alertness	Block	-0.18	0.03	-5.98	-0.70	> 100,000
	Condition (time)	0.41	0.20	2.06	0.49	1.61
	Block x Condition (time)	0.05	0.04	1.08	0.13	0.24

Note. DV = dependent variable, B = standardized regression coefficient, SE = standard error of regression coefficient, BF₁₀ = Ratio of evidence (Bayes Factor) in favor of effect.

Goal commitment ratings

In addition the motivation and alertness ratings, participants made a rating about their current level of goal commitment at the end of each block. Specifically, a screen asked them, “How committed are you to meeting the goal?” and they responded on a 7-point scale (1 = “not committed at all”, 7 = “totally committed”). Due to a programming error, these responses were not recorded for the no-incentive condition. But we report the data for the cash- and time-incentive conditions.

Post-experimental questionnaire

See Experiment 1.

Results & Discussion

Because we did not have any *a priori* hypotheses about the differential impact of the two different incentive types, we separately compared each incentive condition to the no-incentive condition. Figure 5 shows data for each dependent variable, and Tables 7 and 8 summarize the linear mixed effects models. First, comparing the time incentive to the no-incentive condition revealed that participants with a time incentive did not differ from participants in the no-incentive condition in their task performance or number of TUTs reported. There were modest effects of the time incentive motivation and alertness, but the

Table 8

Summaries of models comparing cash and no-incentive conditions in Experiment 3

DV	Effect	B	SE	<i>t</i>	<i>d</i>	BF ₁₀
Reaction times (quintile)	Bin	0.38	0.01	52.51	1.05	> 100,000
	Condition (cash)	-0.06	0.11	-0.51	-0.11	1.63
	Bin x Condition (cash)	0.05	0.01	4.52	0.09	0.23
Reaction times (block)	Block	0.07	0.01	8.14	0.16	> 100,000
	Condition (cash)	0.08	0.10	0.76	0.18	1.63
	Block x Condition (cash)	0.02	0.01	1.65	0.03	0.23
Task-unrelated thoughts	Block	0.05	0.01	4.60	0.24	> 100,000
	Condition (cash)	-0.04	0.05	-0.80	-0.19	0.20
	Block x Condition (cash)	0.00	0.02	0.20	0.01	0.10
Motivation	Block	-0.19	0.03	-6.99	-0.81	> 100,000
	Condition (cash)	-0.17	0.19	-0.86	-0.20	0.41
	Block x Condition (cash)	-0.02	0.04	-0.54	-0.06	0.11
Alertness	Block	-0.18	0.03	-5.95	-0.69	> 100,000
	Condition (cash)	-0.16	0.18	-0.89	-0.21	0.38
	Block x Condition (cash)	-0.04	0.04	-1.04	-0.12	0.18

Note. DV = dependent variable, B = standardized regression coefficient, SE = standard error of regression coefficient, BF₁₀ = Ratio of evidence (Bayes Factor) in favor of effect.

Table 9

Correlations among dependent variables in Experiment

	1	2	3	4	5
1. Reaction time	–				
2. Task-unrelated thoughts	.50	–			
3. Motivation	-.46	-.51	–		
4. Alertness	-.48			–	
5. Goal commitment					.70

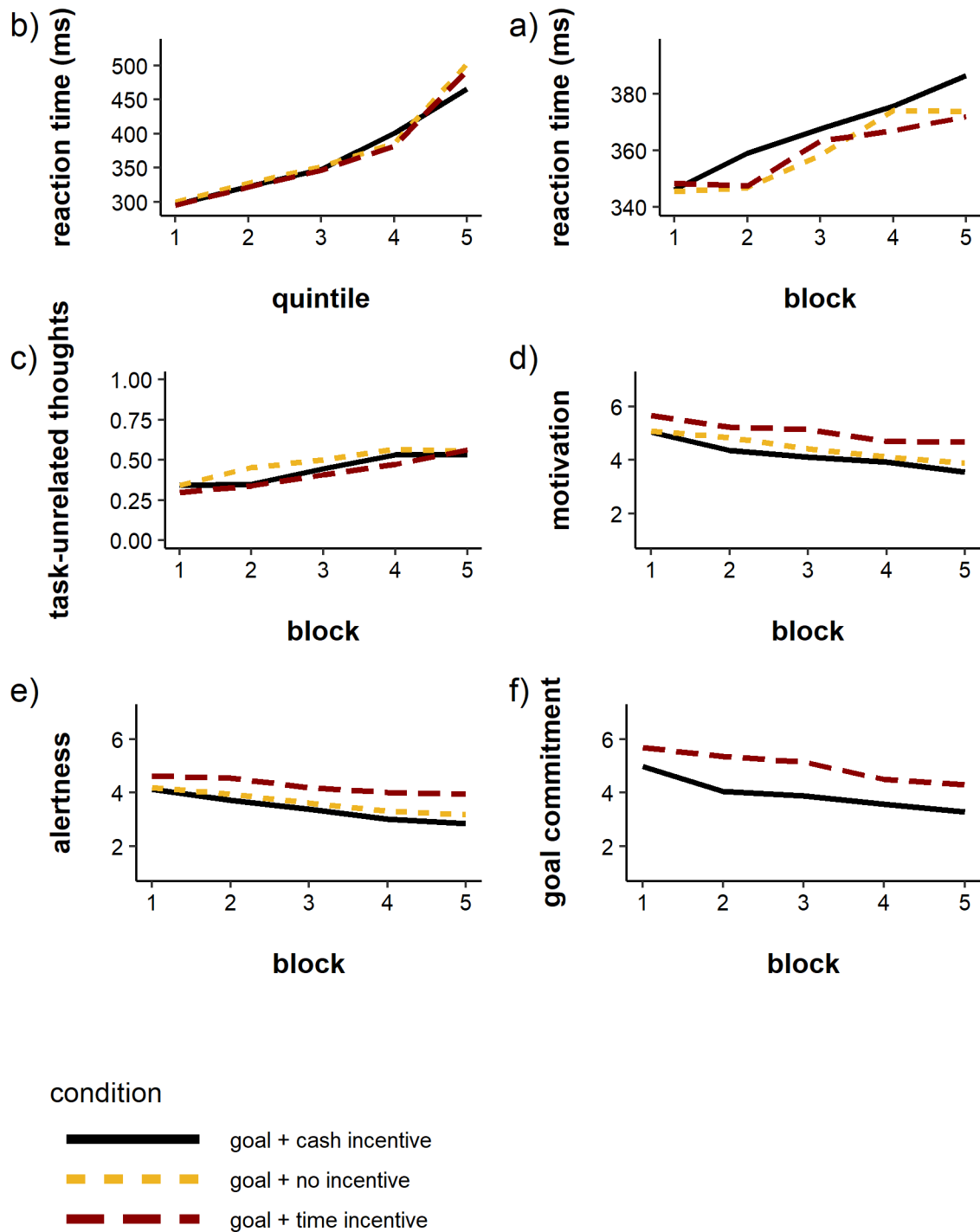


Figure 5. a) Reaction times, b) task-unrelated thoughts, c) motivation ratings, and d) alertness ratings by block and condition in Experiment 3.

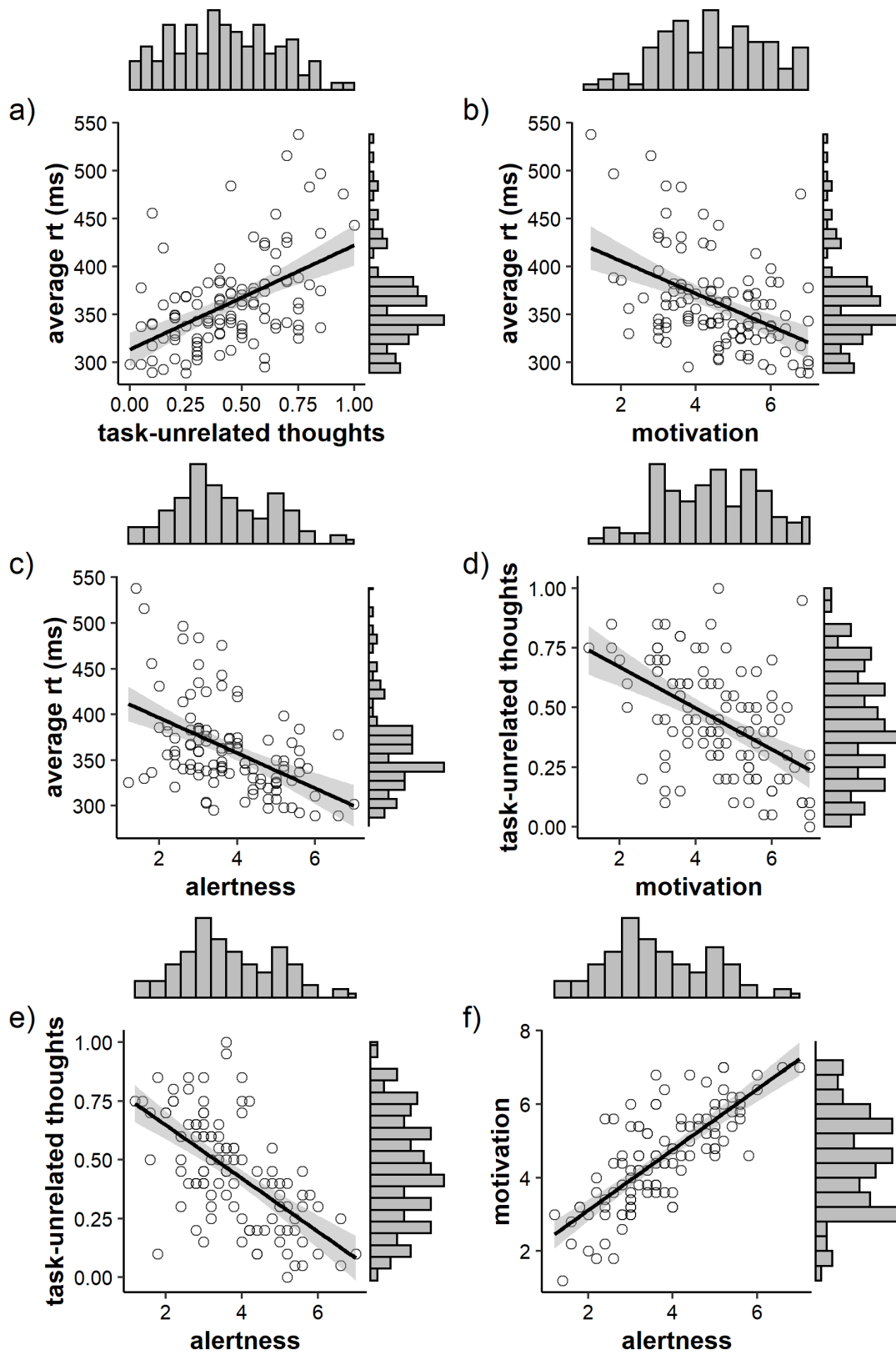


Figure 6. Scatterplots of correlations among dependent variables in Experiment 3.

BFs did not strongly favor inclusion of the effect (BFs ~ 1.6). So although offering early release from the experiment did slightly alter how motivated and alert people reported feeling, it certainly did not actually change their task performance. Next, comparing dependent variables between the cash incentive condition and the no-incentive condition, RTs were very slightly faster in the cash condition compared to the no-incentive condition, but these effects were small and the data were indiscriminate between inclusion and exclusion of the effect of incentive (BF ~ 1.6).

Due to a programming error, we only had goal commitment response data from the two incentive conditions. But as is visible in Figure 5f, the two conditions differed in the levels of goal commitment they reported. To examine this statistically, we submitted goal commitment responses to a mixed model with fixed effects of condition (time vs. cash) and block, and a random effect of participant. This model revealed a large effect of block ($B = -0.20$, $SE = 0.02$, $d = -0.97$, $BF_{10} > 100,000$) and a large effect of condition ($B = 0.56$, $SE = 0.19$, $d = 0.68$, $BF_{10} = 8.55$), but no block x condition interaction ($B = 0.01$, $SE = 0.04$, $d = 0.04$, $BF_{10} = 0.10$). Both groups expressed large downward changes in goal commitment across the course of the task. And overall, participants in the time-incentive condition expressed much higher goal commitment than participants in the cash-incentive condition. Finally, we computed inter-individual correlations to ensure the task engagement measures correlated meaningfully with the task-performance measure. Indeed, they did (all $|r|s > .40$). These correlations are listed in Table 9 and depicted visually in Figure 6.

There are at least two potential explanations for why the incentives did not produce better performance in Experiment 3. First, it could be the case that goals and rewards are motivational, but that feedback is a more powerful motivator. In both Experiment 3 and Experiment 4, we included block-by-block feedback in all conditions. Thus, it could be the case that rewards and/or goal difficulty would affect performance in the absence of feedback, but that feedback mitigates the power of these effects when all conditions receive it.⁴ Additionally, the incentive was offered in an all-or-nothing manner. Participants were told if they met the goal, they would receive the reward, and if they did not, they would not receive the reward (or have to complete an additional 30 min of the experiment). There is some research showing that difficult goal work better when paired with piece-rate rewards and that easier goals work better with all-or-nothing rewards (Mowen, Middlemist, & Luther, 1981). Thus, it could have been the case the incentives were not effective because they were paired with a goal that was very difficult to reach.

Experiment 4

In the final experiment, we examined one lingering issue that pervaded the first three experiments. We set a rather difficult standard for participants to meet (keep average RT < 300 ms) in the hard-goal condition in Experiment, the goal conditions in Experiment 2, and all three conditions in Experiment 3. Indeed, very few participants actually met this goal. In Experiment 1, only 1 participant met the difficult goal. In Experiment 2, only 1 participant met the goal. And in Experiment 3, only 7 participants met the goal. Although Locke and

⁴We would like to thank an anonymous reviewer for pointing out this potential explanation.

colleagues argue that performance tends to scale monotonically with goal-difficulty (Locke, Cartledge, & Knerr, 1970), some have argued that optimal motivation to reach a standard is achieved at moderate probabilities of success (e.g., Atkinson, 1957). The fact that we did not observe effects in some conditions may have been that probability of success (i.e., keeping average RT below 300 ms) was perceived as quite low by many participants. In these situations, participants can do one of two things: 1) continually increase effort until there is a minimal discrepancy between their performance and the goal standard, or 2) re-adjust their goal standard to one other than the experimentally-instructed one. Indeed, many people indicated they had done this in the post-experimental questionnaires (see General Discussion). Therefore, in Experiment 4 we parametrically manipulated goal-difficulty by including several different goal standards. If the goal we set was indeed *too* difficult in the previous experiments, then we should see maximal performance and motivation and minimal TUTs at moderately-difficult goals.

Method

Participants and procedure

A sample of 183 participants from the human subject pool at Arizona State University completed the study in exchange for partial course credit. Participants first completed an informed-consent form. Then they were randomly assigned to one of five conditions: no-goal, 300-goal, 375-goal, 450-goal, and 525-goal. After excluding 27 participants, the final sample included 156 participants (31 in no-goal condition, 28 in 300-goal condition, 31 in 375-goal condition, 31 in 450-goal condition, 35 in 525-goal condition; 68 women, 86 men, 2 other/non-binary gender, $M_{age} = 19.26$, $SD_{age} = 19.26$).

Psychomotor vigilance task

The task was identical to that used in the feedback + goal condition of Experiment 2 with the exception that the goal was manipulated across conditions.

Goal-setting

The experiment included five between-subjects conditions: no-goal, 300-goal, 375-goal, 450-goal, 525-goal. In the no-goal condition, participants were instructed, “Previous research has demonstrated that good performance on this task is a fast reaction time. So your goal should be to keep your average reaction time as low as possible.” In the 300-goal condition, participants were instructed, “Previous research has demonstrated that good performance on this task is a reaction time below 300 milliseconds. So your goal should be to keep your average reaction time below 300 milliseconds.” In the 375-goal condition, participants were instructed, “Previous research has demonstrated that good performance on this task is a reaction time below 375 ms. So your goal should be to keep your average reaction time below 375 milliseconds.” In the 450-goal condition, participants were instructed, “Previous research has demonstrated that good performance on this task

Table 10
Summaries of models with linear goal-difficulty contrast in Experiment 4

DV	Effect	B	SE	<i>t</i>	<i>d</i>	BF ₁₀
Reaction times (quintile)	Bin	0.40	0.01	60.33	0.92	> 100,000
	Goal difficulty (linear)	-0.01	0.03	-0.44	-0.07	0.07
	Bin x Goal difficulty	0.01	0.00	1.39	0.02	0.18
Reaction times (block)	Block	0.11	0.01	13.79	0.21	> 100,000
	Goal difficulty (linear)	0.00	0.03	0.06	0.01	0.11
	Block x Goal difficulty	-0.01	0.00	-1.55	-0.02	0.10
Task-unrelated thoughts	Block	0.09	0.01	8.50	0.35	> 100,000
	Goal difficulty (linear)	-0.04	0.02	-2.14	-0.39	1.14
	Block x Goal difficulty	-0.02	0.01	-3.15	-0.13	10.66
Motivation	Block	-0.23	0.03	-8.86	-0.79	> 100,000
	Goal difficulty (linear)	-0.01	0.07	-0.16	-0.03	0.27
	Block x Goal difficulty	0.03	0.01	1.80	0.16	0.50
Alertness	Block	-0.26	0.03	-8.45	-0.76	> 100,000
	Goal difficulty (linear)	-0.07	0.06	-1.20	-0.22	0.44
	Block x Goal difficulty	0.05	0.02	2.87	0.26	6.24

Note. DV = dependent variable, B = standardized regression coefficient, SE = standard error of regression coefficient, BF₁₀ = Ratio of evidence (Bayes Factor) in favor of effect.

is a reaction time below 450 ms. So your goal should be to keep your average reaction time below 450 milliseconds.” In the 525-goal condition, participants were instructed, “Previous research has demonstrated that good performance on this task is a reaction time below 525 ms. So your goal should be to keep your average reaction time below 525 milliseconds.”

Thought probes

See Experiment 1.

Motivation and alertness ratings

See Experiment 1.

Results & Discussion

The data are plotted in Figure 7. Our first analysis tested the hypothesis that there is a linear relationship between goal difficulty, task performance, and task engagement. To perform this analysis, we created a linear contrast in the data with increasing goal difficulties (525 - 450 - 375 - 300). This was entered as a fixed effect in the models with block as the other fixed factor. The models are summarized in Table 10. Goal difficulty did not have an effect on RTs or motivation, but there was fairly strong evidence that the TUT scaled

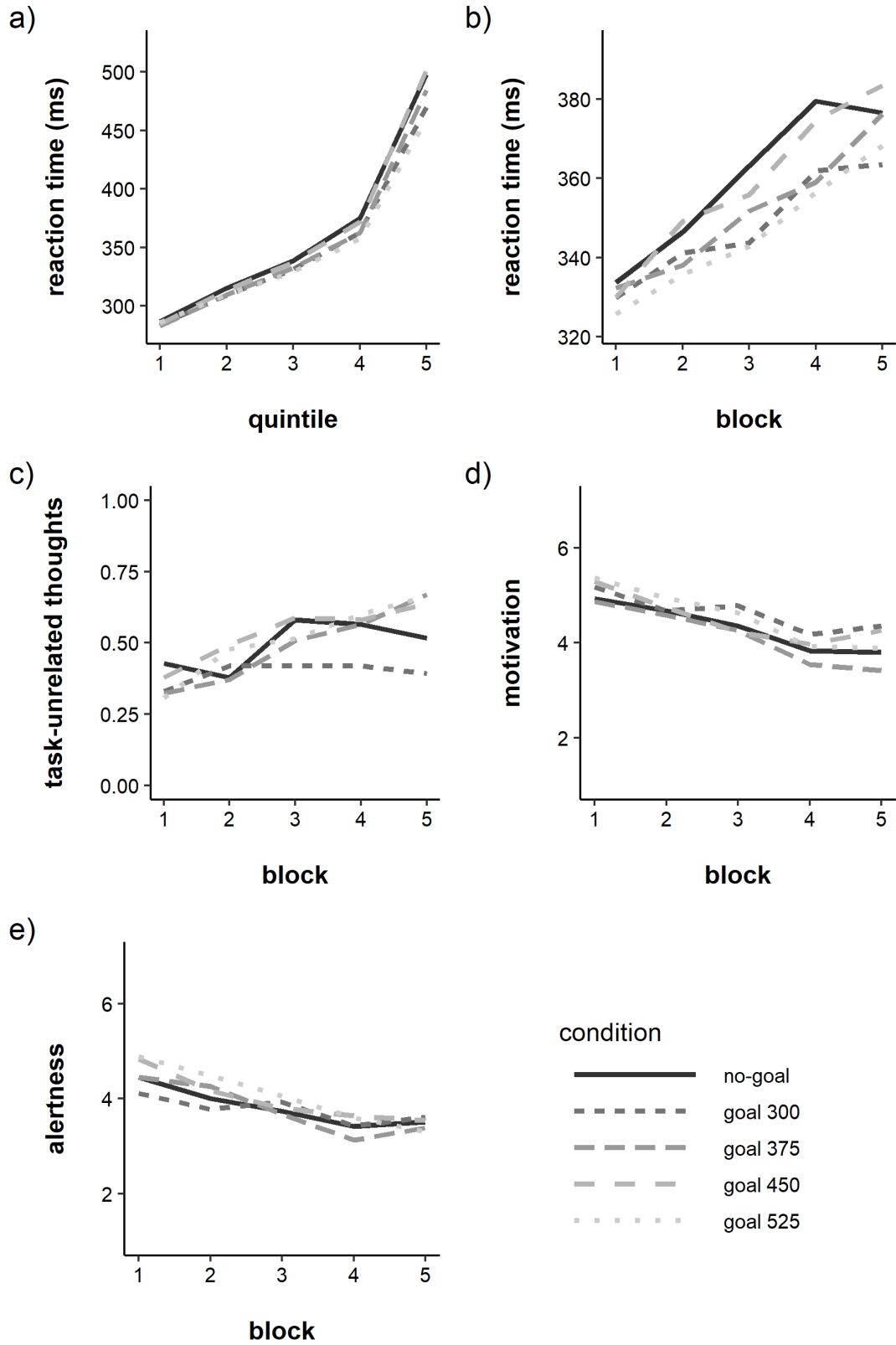


Figure 7. a) Reaction times, b) task-unrelated thoughts, c) motivation ratings, and d) alertness ratings by block and condition in Experiment 4.

Table 11
Summaries of models with moderate goal-difficulty contrast in Experiment 4

DV	Effect	B	SE	<i>t</i>	<i>d</i>	BF ₁₀
Reaction times (quintile)	Bin	0.37	0.01	65.08	1.00	> 100,000
	Goal difficulty (moderate)	-0.10	0.07	-1.59	-0.25	0.16
	Bin x Goal difficulty	0.06	0.01	7.55	0.12	> 100,000
Reaction times (block)	Block	0.09	0.01	12.80	0.20	> 100,000
	Goal difficulty (moderate)	0.08	0.06	1.30	0.23	0.16
	Block x Goal difficulty	0.02	0.01	2.29	0.04	0.43
Task-unrelated thoughts	Block	0.05	0.01	5.67	0.24	> 100,000
	Goal difficulty (moderate)	0.05	0.04	1.23	0.22	0.28
	Block x Goal difficulty	0.02	0.01	1.74	0.07	0.35
Motivation	Block	-0.19	0.02	-8.24	-0.74	> 100,000
	Goal difficulty (moderate)	-0.16	0.15	-1.06	-0.19	0.42
	Block x Goal difficulty	-0.01	0.03	-0.39	-0.03	0.12
Alertness	Block	-0.18	0.03	-6.66	-0.60	> 100,000
	Goal difficulty (moderate)	-0.03	0.13	-0.22	-0.04	0.21
	Block x Goal difficulty	-0.02	0.04	-0.55	-0.05	0.15

Note. DV = dependent variable, B = standardized regression coefficient, SE = standard error of regression coefficient, BF₁₀ = Ratio of evidence (Bayes Factor) in favor of effect.

Table 12
(ref:e4-corr-caption)

	1	2	3	4
1. Reaction time	–			
2. Task-unrelated thoughts	.22	–		
3. Motivation	-.17	-.53	–	
4. Alertness	-.22	-.50	.69	–

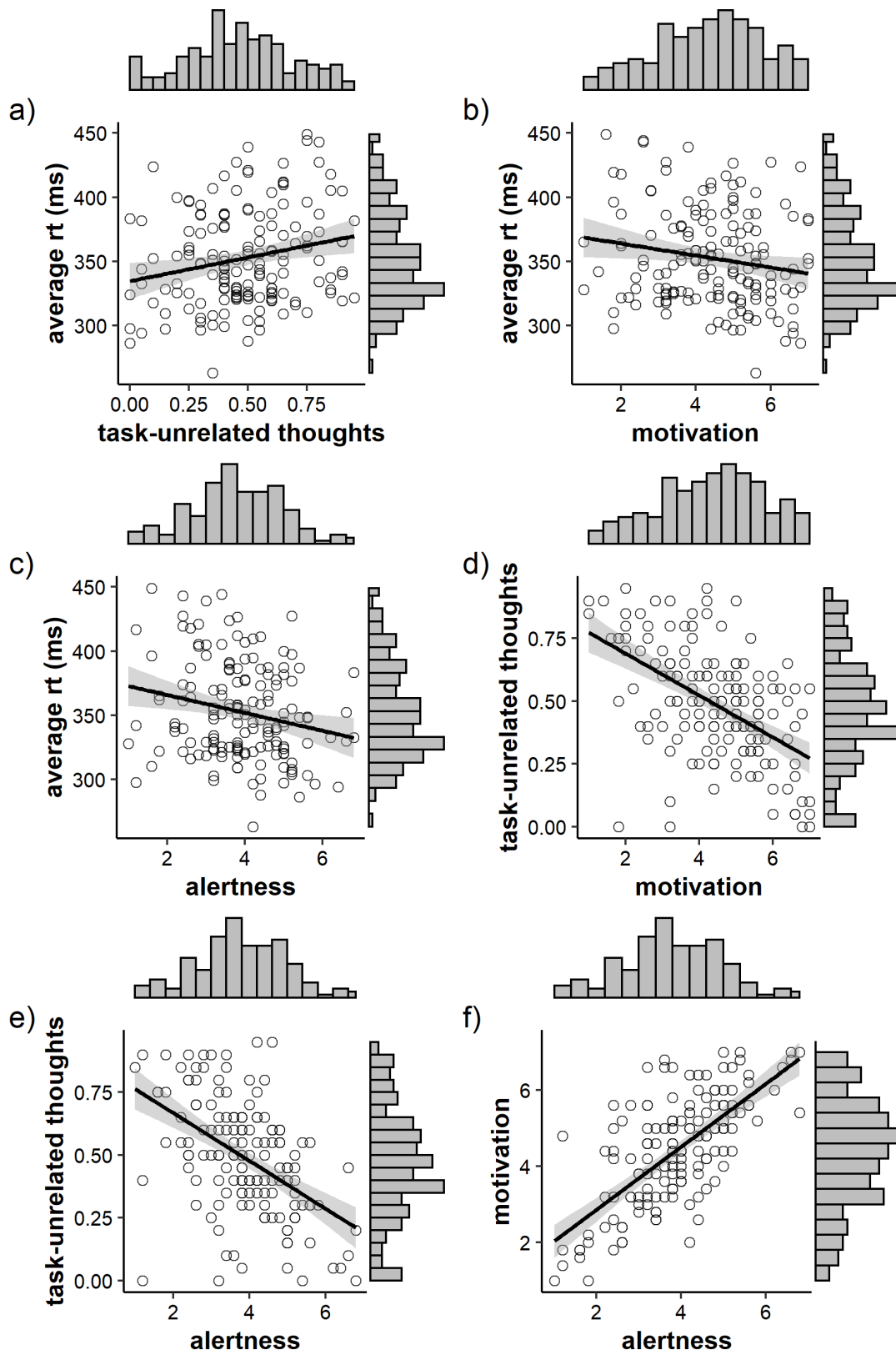


Figure 8. Scatterplots of correlations among dependent variables in Experiment 4.

negatively with goal difficulty, and that the change in TUT across time was lower at higher goal difficulties (see Figure 7c). There was a small block x goal difficulty interaction on alertness. The easier goal conditions showed steeper declines in alertness across blocks.

The second analysis tested the hypothesis that optimal task performance and task engagement would be achieved at moderately-difficult goals. To perform this analysis, we created a contrast in the data for which the two moderately-difficult goals (375 and 450) were compared to the easiest and most difficult goals (525 and 300). These models are summarized in Table 11. For the most part, there was no strong evidence for any dependent variable that matched predictions made by this hypothesis. In some cases, the data were indiscriminate between models with and without an effect of moderate goal difficult. There was a strong evidence for a bin x moderate goal interaction, but in the opposite direction of what would be hypothesized. Participants in the moderately-difficult goal conditions had the longest RTs at the slowest end of the distribution. But overall, there was not a pattern of evidence consistent with the hypothesis that moderately-difficult goals lead to optimal sustained attention. Interindividual correlations among the dependent variables were lower than in Experiments 1, 2, and 3 (see Table 12 and Figure 8).

General Discussion

Sustained attention/vigilance tasks are characterized by decrements in performance across time. In the psychomotor vigilance task, this manifests as a slowing of RTs across time (Brewer et al., 2017; Massar et al., 2016, 2019; Unsworth & Robison, 2016, 2020). The slowing of RTs is accompanied by decreases in self-reports of motivation, decreases in physiological arousal (Unsworth & Robison, 2016), and an increase in task-unrelated thoughts (TUTs; Unsworth & Robison, 2016). A major question is whether vigilance decrements can be mitigated or eliminated. Several attempts have been successful, including providing monetary incentives (Esterman et al., 2016; Massar et al., 2019) and providing feedback (Sipowicz et al., 1962). However several studies have failed to find mitigation or elimination of the vigilance decrement, despite main effects of experimental manipulations on *overall* task performance (Brewer et al., 2017; Esterman et al., 2014; Massar et al., 2016; Warm et al., 1973). Here, we leveraged theories from industrial-organizational psychology to systematically examine three potential ways to improve sustained attention (goal-setting, feedback, and incentives) in hopes of attenuating or eliminating vigilance decrements. Another major goal of the present study was to collect multimodal data within the context of a sustained attention task. This allowed us to examine the subjective as well and objective signatures of the vigilance decrement. These data would also allow us to test predictions made by various theories of vigilance. Thus, the results have implications for goal-setting theory and for theoretical accounts of the vigilance decrement.

Theoretical implications

The present study has implications for two sets of theories: goal-setting theory and theories of vigilance. We based our hypotheses on goal-setting theory, as well as the moderators proposed by Locke and Latham (2002). In Experiment 1, we hypothesized that setting

a specific, difficult goal for participants would speed RTs, attenuate the vigilance decrement, increase motivation, and decrease task-unrelated thoughts. The data indicated that participants who were given a specific goal, regardless of difficulty, demonstrated faster RTs and an attenuated vigilance decrement. More specifically, the main effect on RTs was qualified by a bin x goal-specificity interaction. The specific goals mostly reduced the longest RTs in the distribution. Interestingly, these effects were not accompanied by increased motivation or decreased TUTs. Additionally, the difficult goal did not affect the dependent variables over and above an easy yet specific goal. The only effect for which there was some evidence was a bin x goal-difficulty interaction, wherein participants with the more difficult goal tended to show faster RTs at the fast end of the distribution. Thus, there was evidence that goal *specificity* could reduce the vigilance decrement, but not necessarily increase task engagement. And for the most part, there was very limited evidence for the hypothesis that a difficulty goal would be more effective than an easy goal.

In Experiment 2, we tested the effectiveness of feedback on modulating sustained attention. To do so, we crossed goal-setting and feedback to examine their unique and interactive effects. Based on goal-setting theory, we hypothesized that goal-setting and feedback would interact such that the fastest RTs, the shallowest vigilance decrements, the highest motivation, and the fewest TUTs would be produced by the combination of a specific, difficult goal and performance feedback. The data supported several of these hypotheses: there was strong evidence for a block x goal interaction, replicating Experiment 1, and a block x goal x feedback interaction, indicating that the goal + feedback condition had the shallowest vigilance decrement. Finally, there was strong evidence for a bin x goal x feedback interaction. The largest effect of the goal x feedback combination was on the slowest RTs of the distribution. Feedback also significantly reduced TUTs, mostly due to a decrease in *intentional* mind-wandering. Collectively, these findings largely support our hypothesis that goal-setting and feedback can produce better task performance and higher task engagement. In Experiment 3, we hypothesized that pairing a difficult goal with an incentive would speed RTs, attenuate the vigilance decrement, increase motivation, and decrease TUTs. We also hypothesized that providing an incentive would increase goal commitment. Interestingly, the incentives had few effects on our dependent measures. Offering a time-based incentive (i.e., early release from the experiment) only increased motivation and alertness very slightly, and the evidence for this effect was not strong. Further, the time incentive had no effect on task performance. The cash incentive did not have an effect on any of the task-engagement measures, and the data did not strongly favor its effect on task performance. In Experiment 4, we examined the possibility that our difficult goal may have been *too* difficult for participants, and thus served as a disincentivizing force rather than a motivational one. To do so, we specified a variety of goals. If optimal goal pursuit is attained at moderately-difficult goal levels (Atkinson, 1957), we should have observed the fastest RTs, the shallowest vigilance decrement, the highest motivation, and the fewest TUTs in the moderate goal conditions. However this was not the case. So our null findings in Experiments 1, 2, and 3 were probably not due to our difficult goal being too difficult to attain.

One interesting aspect of the data were the individual differences in goal coherence. As a manipulation check, we asked every participant in Experiment 1, 2, and 3 to give an open-

ended report of their goal at the end of the task. We received a wide variety of responses to this question. Many participants did indeed report the goal that corresponded to their condition. However, many participants reported adjusting their goals. In Experiment 1, only 9 of 35 participants (26%) in the hard-goal condition mentioned the goal of keeping their reaction time below 300 ms, and only one participant in the easy-goal condition specifically mentioned trying to keep their RTs below 800 ms. In Experiment 2, only 6% of participants in the goal + no-feedback condition mentioned the 300-ms goal in their post-experiment questionnaire, and only 47% of participants in the goal + feedback condition mentioned the 300-ms goal in their response. In Experiment 3, a larger proportion of respondents mentioned the goal: 59% in the no-incentive condition, 75% in the time incentive condition, and 68% in the cash incentive condition. The responses did map on to patterns of task performance, however. For example, one participant said, “At first I wanted to see how fast I could react to the changing numbers. I wanted to be as fast as I could. That desire waned as the task progressed and I got tired/sore.” Another participant said, “[My goal was] to do as instructed to the best of my ability, pressing the button as soon as possible. My resolve to accomplish this goal gradually ebbed.” Finally, there were large decreases in goal commitment ratings across time in Experiment 3. This indicates that participants do not always maintain a high level of commitment to an externally-directed goal. In ongoing work, we are following up on this finding by allowing participants to set their own goals and seeing how commitment to such goals changes across time.

A second major theoretical question addressed in the current study is whether vigilance decrements can be mitigated or eliminated via various intervention techniques. One major theory of the vigilance decrement is that vigilance/sustained attention situations deplete a limited resource (Warm, Dember, & Hancock, 1996; Warm et al., 2008). Performance decrements are thus observed because of this resource depletion. However others have argued that the fact that performance can be “restored” with certain interventions, even without any form of rest or replenishment, is inconsistent with this supposition (Brewer et al., 2017; Hopstaken et al., 2015b). Rather, the effects may be due to a shift toward task disengagement because the task becomes less and less rewarding across time (Boksem & Tops, 2008; Kurzban et al., 2013; Thomson et al., 2015). Here, we tried to make the tasks more rewarding in various ways. Two of our manipulations did indeed attenuate the vigilance decrement: setting a specific goal, even further providing performance feedback as to how participants were meeting said goal. Thus, it appears that there are ways to offset the vigilance decrement in this task.

Our results can be interpreted in one of two ways. One way to interpret our results would be that they fall in line with predictions made by the resource account. It is clear that participants were experiencing mental fatigue over the course of the task. In addition to showing slower reaction times, participants also reported more task-unrelated thoughts, decreasing motivation, and decreasing alertness. Further, even conditions that reduced the vigilance decrement did not eliminate it. Thus the results could be interpreted to mean that the vigilance decrement is something obligatory - it happens no matter how we motivate people. However, many people expressed decreasing willingness to engage with the task and a shifting of goals as time progressed, as evidenced by their self-reports of motivation and the post-experimental questionnaires. That suggests at least a certain degree of task

disengagement is something that participants are doing on purpose, rather than something that is happening to them as a result of resource depletion. This finding is consistent with models that argue that the vigilance decrement is at least partially due to a shifting of attention away from the task and to other thoughts (e.g., mind-wandering; Thomson et al., 2015; Boksem & Tops, 2008; Kurzman et al., 2013). So it is possible that both aspects of vigilance are at play: there is resource depletion occurring, resulting in a degree of mental fatigue demonstrated by participants in all conditions. But, perhaps depletion can be offset, at least to a certain degree, by an individual's exertion of additional effort to reach a particular goal.

It could also have been the case that our reward manipulations were not sufficiently motivational. In other studies, Massar et al. (2016) offered a monetary reward in a piece-rate manner. Specifically, Massar et al. had participants complete one baseline run of the PVT, then two incentivized runs. In the low-reward run, participants were given 1 cent for every fast response, and in the high-reward run, they were given 10 cents for every fast response. Fast responses were categorized anything faster than a participant's median reaction time during the baseline run. Massar et al. showed that participants responded significantly faster in the low-reward condition compared to baseline, and significantly faster in the high-reward condition compared to the low-reward condition. However, there was still a vigilance decrement in all three runs, and there was no difference in the magnitude of the vigilance decrement across the three runs. In a second study, Massar et al. (2019) compared various incentive structures in sleep-deprived versus rested participants in the PVT. The rewards were 1, 5, and 10 cents for each fast trial. In this study, both sleep deprived and rested participants showed shallower vigilance decrements with higher rewards. The crucial difference between Massar et al.'s studies and the present study is a piece-rate reward structure vs. an all-or-nothing reward structure. In their study examining motivational manipulations in the PVT, Brewer et al. (2017)'s incentive decreased RTs, and this effect was largest early on during a second run of the PVT. But it did not mitigate the vigilance decrement. We did not observe a main effect of either incentive on RTs. Perhaps piece-rate incentive structures and competitive situations are more motivating than the all-or-nothing reward we offered in Experiment 3. Finally, it is possible that there is an interaction between incentive structure and goal difficulty. For example, Mowen et al. (1981) gave participants one of two goals (easy or difficult) and one of two incentive structures (piece-rate or all-or-nothing). The data showed a crossover interaction. Under piece-rate incentive structures where the reward scaled with proximity to the goal standard. But under an all-or-nothing incentive structure in which the reward was only earned if participants met or exceeded the goal, perform a hard goal produced better performance than an easy goal. Therefore in ongoing work, we are exploring piece-rate incentive structures to modulate people's willingness to engage and sustain their attention. Future work is still necessary to test predictions made by the various models of sustained attention described above.

Conclusion

We took a theory-driven experimental approach motivated by predictions made by goal-setting theory. We found evidence that setting a specific goal improves sustained at-

tention via a reduction in the vigilance decrement and a shortening of the longest response times, albeit with no effect on task engagement, that feedback increases both task performance and task engagement, and that pairing goals with feedback is particularly effective in increasing task engagement. We found very little evidence that incentives (a cash bonus or early release from the experiment) can increase task engagement or task performance beyond the effects of goals and/or feedback. In future research, we plan to leverage these findings to more specifically test predictions made by competing theories of vigilance.

References

- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*, 359–372.
- Auguie, B. (2017). *GridExtra: Miscellaneous functions for "grid" graphics*. Retrieved from <https://CRAN.R-project.org/package=gridExtra>
- Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown*. Retrieved from <https://github.com/crsh/papaja>
- Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science*, *20*(3), 160–166. <https://doi.org/10.1177/0963721411408885>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bills, A. G. (1931a). Blocking: A new principle of mental fatigue. *The American Journal of Psychology*, *43*, 230–245.
- Bills, A. G. (1931b). Mental work. *Psychological Bulletin*, *28*, 505–532.
- Boksem, M. A., & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews*, *59*, 125–139.
- Brewer, G. A., Lau, K. K., Wingert, K. M., Ball, B. H., & Blais, C. (2017). Examining depletion theories under conditions of within-task transfer. *Journal of Experimental Psychology: General*, *146*, 988–1008.
- Church, R. M., & Camp, D. S. (1965). Change in reaction-time as a function of knowledge of results. *The American Journal of Psychology*, *78*, 102–106.
- Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual rt task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, *17*, 652–655.
- Drummond, S. P., Bischoff-Grethe, A., Dinges, D. F., Ayalon, L., Mednick, S. C., & Meloy, M. (2005). The neural basis of the psychomotor vigilance task. *Sleep*, *28*, 1059–1068.
- Erez, M. (1977). Feedback: A necessary condition for the goal setting-performance relationship. *Journal of Applied Psychology*, *62*(5).
- Esterman, M., Grosso, M., Liu, G., Mitko, A., Morris, R., & DeGutis, J. (2016). Anticipation of monetary reward can attenuate the vigilance decrement. *PloS One*, *11*, e0159741.
- Esterman, M., Reagan, A., Liu, G., Turner, C., & DeGutis, J. (2014). Reward reveals dissociable aspects of sustained attention. *Journal of Experimental Psychology: General*, *143*, 2287–2295.
- Esterman, M., & Rothlein, D. (2019). Models of sustained attention. *Current Opinion in Psychology*, *29*, 174–180.

- Feng, S., D’Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review*, *20*, 586–592.
- Hopstaken, J. F., Linden, D. van der, Bakker, A. B., & Kompier, M. A. (2015a). The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics. *Biological Psychology*, *110*, 100–106.
- Hopstaken, J. F., Van Der Linden, D., Bakker, A. B., & Kompier, M. A. (2015b). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology*, *52*, 305–315.
- Kleiman, E. (2017). *EMAtools: Data management tools for real-time monitoring/ecological momentary assessment data*. Retrieved from <https://CRAN.R-project.org/package=EMAtools>
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*, 661–679.
- Latham, G. P., & Kinne, S. B. (1974). Improving job performance through training in goal setting. *Journal of Applied Psychology*, *59*, 187.
- Locke, E. A., & Bryan, J. F. (1969). Knowledge of score and goal level as determinants of work rate. *Journal of Applied Psychology*, *53*(1).
- Locke, E. A., Cartledge, N., & Knerr, C. S. (1970). Studies of the relationship between satisfaction, goal-setting, and performance. *Organizational Behavior and Human Performance*, *5*(2).
- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Prentice-Hall, Inc.
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*, 705.
- Locke, E. A., & Latham, G. P. (2019). The development of goal setting theory: A half century retrospective. *Motivation Science*, *5*, 93–105.
- Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin*, *90*, 125–152.
- Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*, 6–21.
- MacLean, K. A., Ferrer, E., Aichele, S. R., Bridwell, D. A., Zanesco, A. P., Jacobs, T. L., . . . others. (2010). Intensive meditation training improves perceptual discrimination and sustained attention. *Psychological Science*, *21*, 829–839.
- Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia*, *37*, 661–670.
- Massar, S. A., Lim, J., Sasmita, K., & Chee, M. W. (2016). Rewards boost sustained attention through higher effort: A value-based decision making approach. *Biological Psychology*, *120*, 21–27. <https://doi.org/10.1016/j.biopsycho.2016.07.019>

- Massar, S. A., Lim, J., Sasmita, K., & Chee, M. W. (2019). Sleep deprivation increases the costs of attentional effort: Performance, preference and pupil size. *Neuropsychologia*, *123*, 169–177.
- McCormack, P. (1959). Performance in a vigilance task with and without knowledge of results. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *13*, 68–71.
- McCormack, P., Binding, F., & Chylinski, J. (1962). Effects on reaction-time of knowledge of results of performance. *Perceptual and Motor Skills*, *14*, 367–372.
- McCormack, P., Binding, F., & McElheran, W. (1963). Effects on reaction time of partial knowledge of results of performance. *Perceptual and Motor Skills*, *17*, 279–281.
- McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 196.
- McVay, J. C., & Kane, M. J. (2012). Drifting from slow to “d’oh!”: Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 525–549.
- Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of bayes factors for common designs*. Retrieved from <https://CRAN.R-project.org/package=BayesFactor>
- Mowen, J. C., Middlemist, R. D., & Luther, D. (1981). Joint effects of assigned goal level and incentive structure on task performance: A laboratory study. *Journal of Applied Psychology*, *66*, 598–603.
- Payne, R., & Hauty, G. (1955). Effect of psychological feedback upon work decrement. *Journal of Experimental Psychology*, *50*, 343–351.
- Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*, 747–758.
- Robison, M. K. (2018). Regulating mind-wandering and sustained attention with goal-setting, feedback, and incentives. *Doctoral Dissertation*. Retrieved from <http://hdl.handle.net/1794/23712>
- Robison, M. K., Miller, A. L., & Unsworth, N. (2019). Examining the effects of probe frequency, response options, and framing within the thought-probe method. *Behavior Research Methods*, *51*, 398–408.
- Robison, M. K., & Unsworth, N. (2019). Pupillometry tracks fluctuations in working memory performance. *Attention, Perception, & Psychophysics*, *81*, 407–419. <https://doi.org/10.3758/s13414-018-1618-4>
- Rummel, J., & Boywitt, C. D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. *Psychonomic Bulletin & Review*, *21*, 1309–1315.

- Seli, P., Cheyne, J. A., & Smilek, D. (2013). Wandering minds and wavering rhythms: Linking mind wandering and behavioral variability. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 1–5.
- Seli, P., Schacter, D. L., Risko, E. F., & Smilek, D. (2019). Increasing participant motivation reduces rates of intentional and unintentional mind wandering. *Psychological Research*, *83*, 1057–1069.
- Sipowicz, R. R., Ware, J. R., & Baker, R. A. (1962). The effects of reward and knowledge of results on the performance of a simple vigilance task. *Journal of Experimental Psychology*, *64*, 58–61.
- Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition*, *13*, 657–690.
- Stawarczyk, D., Majerus, S., Maj, M., Van der Linden, M., & D'Argembeau, A. (2011). Mind-wandering: Phenomenology and function as assessed with a novel experience sampling method. *Acta Psychologica*, *136*, 370–381.
- Strang, H. R., Lawrence, E. C., & Fowler, P. C. (1978). Effects of assigned goal level and knowledge of results on arithmetic computation: A laboratory study. *Journal of Applied Psychology*, *63*, 446–450.
- Thomson, D. R., Besner, D., & Smilek, D. (2013). In pursuit of off-task thought: Mind wandering-performance trade-offs while reading aloud and color naming. *Frontiers in Psychology*, *4*, 360.
- Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Perspectives on Psychological Science*, *10*.
- Thorndike, E. L. (1912). The curve of work and the curve of satisfyingness. *Journal of Applied Psychology*, *1*, 165–194.
- Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience*, *16*, 601–615. <https://doi.org/10.3758/s13415-016-0417-4>
- Unsworth, N., & Robison, M. K. (2018). Tracking arousal state and mind wandering with pupillometry. *Cognitive, Affective, & Behavioral Neuroscience*, *18*, 638–664. <https://doi.org/10.3758/s13415-018-0594-4>
- Unsworth, N., & Robison, M. K. (2020). Working memory capacity and sustained attention: A cognitive-energetic perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*, 77–103.
- Unsworth, N., Robison, M. K., & Miller, A. L. (2018). Pupillary correlates of fluctuations in sustained attention. *Journal of Cognitive Neuroscience*, *30*, 1241–1253. https://doi.org/10.1162/jocn_a_01251

- Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 183–200). Erlbaum.
- Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, *50*, 433–441.
- Warm, J. S., Riechmann, S. W., Grasha, A. F., & Seibel, B. (1973). Motivation in vigilance: A test of the goal-setting hypothesis of the effectiveness of knowledge of results. *Bulletin of the Psychonomic Society*, *1*, 291–292.
- Weinstein, Y. (2018). Mind-wandering, how do i measure thee with probes? Let me count the ways. *Behavior Research Methods*, *50*, 642–661.
- Wickham, H. (2017). *Tidyverse: Easily install and load the 'tidyverse'*. Retrieved from <https://CRAN.R-project.org/package=tidyverse>
- Xu, J., & Metcalfe, J. (2016). Studying in the region of proximal learning reduces mind wandering. *Memory & Cognition*, *44*(5), 681–695.