Journal of Experimental Psychology:
Learning, Memory, and Cognition
2016, Vol. 42, No. 6, 897–913

0278-7393/16/$12.00 http://dx.doi.org/10.1037/xlm0000215

Do Participants Differ in Their Cognitive Abilities, Task Motivation, or Personality Characteristics as a Function of Time of Participation?

Matthew K. Robison and Nash Unsworth
University of Oregon

Four experiments tested the conventional wisdom in experimental psychology that participants who complete laboratory tasks systematically differ in their cognitive abilities, motivational levels, and personality characteristics as a function of the time at which they participate during an academic term. Across 4 experiments with over 2,900 participants from 2 different universities with 2 different academic schedules, no convincing evidence suggested that individuals differ in cognitive abilities (working memory capacity, fluid intelligence, crystallized intelligence, long-term memory, and attention control). Similarly, no evidence suggested participants' task motivation varies systematically with time of participation, nor do any of the Big Five personality traits. The present study concludes that researchers need not be overly concerned with time of participation effects as a potential confound in individual differences or experimental psychology.

Keywords: time of participation, individual differences, motivation, personality

The bulk of psychological research on healthy adult populations uses undergraduate participant pools as a sampling of the population. As part of a course requirement, students are typically required to participate in experiments for a certain number of hours or complete an alternative assignment (e.g., write summaries of research articles). One major concern with drawing from this sample is the unmeasured effect of various individual differences on the dependent variables of interest. These unmeasured differences may include cognitive abilities, motivation levels, and personality characteristics that could covary with time of participation. So one artifact of collecting from this sample may be that students who participate early in the academic term qualitatively and significantly differ from those who participate at the end of the term. Also, different academic calendars (e.g., 10-week quarters vs. 16-week semesters) may lead to different patterns of participation. For these reasons, some researchers avoid collecting data during certain times of an academic term in order to avoid these confounds. The present study sought to test the conventional wisdom that early participants (i.e., the "punctual") differ from late participants (i.e., the "procrastinators").

Previous studies on cognitive abilities and time of participation have yielded mixed results. Underwood, Schwenn, and Keppel (1964) found no differences in verbal learning performance for subjects drawn at different times in an academic quarter. Page and Lumia (1968) found no effects for time of sampling on performance in a verbal recall task. Wang and Jentsch (1998) did not find any significant differences between early and late volunteers on a paired-associates learning task. Blatt and Quinlan (1967) compared undergraduate students who volunteered during the first week of an academic term with those who volunteered in the last 2 weeks of the term. A random sample from each group revealed no observable differences between groups on the information or vocabulary subscales of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955), college entrance exam scores, grades, or number of extracurricular activities. Yet they did find significant differences between early and late volunteers on the picture arrangement subscale of the WAIS, a death concern questionnaire, future time perspective, and a difference in interference scores on the Stroop Color-Word Test that trended toward significance. On a serial learning task, Richter et al. (1981) found a steeper learning curve for early participants, but overall number of correct was higher for late participants. Finally, Langston, Ohnesorge, Kruley, and Haase (1994) found no evidence for differences in performance in a signal detection task, nor for a text comprehension task as a function of time of participation. Specifically, there was no evidence that early participants were more sensitive, employed more conservative response strategies, worked harder or more diligently, or were more consistent on the task. Participants also showed no differences in the ability to resolve afferents from a text as a function of time of participation.

Another possibility is that students who participate early in an academic term are differentially motivated compared with students who participate at the end of the term, and this affects their performance on laboratory tasks. Nicholls, Loveless, Thomas, Loetscher, and Churches (2015) compared early participants and late participants on a sustained attention to response task (SART). Half of the participants received course credit and half were monetarily compensated for their participation. Nicholls et al. found no difference in SART performance for credit and paid participants early in the term. But paid participants performed better than credit participants late in the term. In the course credit group, late participants showed significantly more response time

This article was published Online First December 7, 2015.

Matthew K. Robison and Nash Unsworth, Department of Psychology, University of Oregon.

Correspondence concerning this article should be addressed to Matthew K. Robison, Department of Psychology, 1227 University of Oregon, Eugene, OR 97403. E-mail: mkr@uoregon.edu

variability than early participants; early and late paid participants did not differ. Finally, measures of motivation (Student Work Performance Inventory; Amabile, Hill, Hennessey, & Tighe, 1994) revealed higher intrinsic motivation for early participants among the credit group and higher extrinsic motivation for paid participants overall. Nicholls et al. (2015) summarize this pattern of results by saying that students who participate early in the term are intrinsically motivated to perform well. Students who participate for course credit late in the term are less enthusiastic about their participation and do less well.

Grimm, Markman, and Maddox (2012) argue that time of participation effects may arise from an interaction between motivational states and task construals. Grimm et al. (2012) posit that students who participate at the end of the term are not unmotivated, but rather have a different motivational state, one that is induced by the time in the semester. At the beginning of the semester students are more likely to be promotion-focused; at the end of the semester students are more likely to be prevention-focused. An experiment revealed an interaction in which students in the beginning of the semester performed better in a gain condition and students at the end of the semester performed better in a loss condition, which would match situational motivational states at those time points. They conclude that students at the beginning of the term are promotion-focused and thus perform better in a gain condition relative to a loss condition because the gain condition matches their motivational state. The opposite is true for prevention-focused students at the end of the semester.

Another potential difference between early participants and late participants is their personality profile. Harber, Zimbardo, and Boyd (2003) found that students with a more future-oriented time perspective participated earlier in the term and more frequently met submission deadlines than present-oriented participants. Stevens and Ash (2001) found a negative correlation between time of participation and the conscientiousness and neuroticism factors of the NEO PI-R (Costa & McRae, 1992) and a positive correlation between openness and week of participation. In other words, early participants tended to be more conscientious and neurotic and late participants tended to be more open to experience. Aviv, Zelenski, Rallo, and Larsen (2002) measured personality characteristics using the NEO PI-R and found negative correlations between week of participation and the anxiety and self-consciousness facets of the neuroticism scale, the straightforwardness and compliance facets of the agreeableness scale, and the order and self-discipline facets of the conscientiousness scale. That is, early participants scored higher on these measures than late participants. They also found positive correlations between week and the warmth, gregariousness, activity, and excitement-seeking facets of the extraversion scale, and the feelings and ideas facets of the openness scale. Late participants scored higher on these measures than early participants. Aviv et al. (2002) note that although the observed effects are small, they are worth consideration if they covary with the dependent variables of interest in an experiment. Witt, Donnellan, and Orlando (2011) measured personality with the IPIP-NEO (Johnson, 2000) and found that the most notable correlations with week were for conscientiousness and its facets of self-efficacy, dutifulness, achievement striving, self-discipline, and cautiousness. Students who participated early in the term tended to have higher levels of conscientiousness. This study utilized both in-person and online surveys for a total sample size of 512 (70% online). In general, the conclusion from these studies is that if any of these personality traits could have an effect on the dependent variables of interest, they should be measured and accounted for in analyses.

While the results for cognitive measures and time of participation effects are mixed, the reviewed literature more firmly suggests there may be some differences between early and late participants in motivation levels and personality characteristics. However, some of the major limitations of these studies include small sample sizes, data collected during only one academic term, use of only one dependent variable for cognitive abilities, and no simultaneous measurement of cognitive abilities, motivation, and personality.

The Present Study

Combining data collected over the course of several years at two public universities—one in the southeastern U.S. and one in the Pacific Northwest—the present study examined whether there is any evidence for systematic differences in cognitive abilities, personality traits, and motivation levels for participants who complete laboratory tasks at different times throughout an academic term. The bulk of the data comes from large individual differences investigations of cognitive abilities. In other individual differences investigations, personality traits and task motivation were measured as well. There are considerably less data for the latter two constructs of interest, but sizable enough sample sizes to investigate any considerable effects of time of participation.

If undergraduate participants differ in their cognitive abilities or task motivation, early participants should outperform late participants who complete the same measures of cognitive abilities. Further if this phenomenon is one that is common across participant pools, it should replicate across universities with different academic schedules (e.g., 16-week semesters and 10-week quarters). It could be that undergraduate participant pools are actually sampling from two different samples with different distributions as far as cognitive abilities: the "punctual" and the "procrastinators." These effects could arise in several ways.

There are several possible explanations for how time of participation effects may arise, and we will attempt to address each one in the following set of experiments. One possibility is that the cumulative effect of the academic term creates a fatigue factor. As the term comes to an end, a variety of sources are competing for undergraduates' attention: final exams, final projects, plans for breaks, and so forth. This leads to a level of fatigue that affects their cognitive performance, which could be related to a "burnout" effect (Kleinsorge, Diestel, Scheil, & Niven, 2014). Another possibility is that undergraduate students are simply less motivated at the end of the term. They are participating as a requirement for a course, and they just want to get the requirement out of the way. This lack of motivation leads to task disengagement and lower scores on cognitive tasks. A third possibility is that early participants have qualitatively different personality characteristics than late participants. They like to get things done early in the term, late participants tend to put things off until the end of the term, and their personality traits reflect these tendencies. Additionally, it could be the case that early participants simply have better cognitive abilities than late participants. On average, they may have better attention control abilities, higher intelligence, and have greater memory capacity. Indeed, a number of studies have found that early participants report higher grade point averages (GPAs) than late participations (e.g., Aviv et al., 2002; Bender, 2007). Finally, it could be the case that some cognitive abilities are more susceptible to time of participation effects than others, as factors like fatigue and task motivation could differentially affect various cognitive abilities. While the present study does not attempt to differentiate various sources of time of participation effects, it tests the hypothesis that early participants differ from late participants in their cognitive abilities, task motivation, and personality traits.

Experiment 1

The first dataset derives from a series of individual differences investigations and experiments collected over a 4-year span at the University of Georgia. The university uses two 16-week academic semesters, each with a 1-week break in the middle of the semester. All participants were granted course credit for participation. Data were rarely collected during the first week of an academic semester as participants were typically unaware of the course requirement for participation until after the first week of classes, and this period of time was used to train research assistants and prepare experimental materials.

Method

Participants. Data were collected from a total of 801 participants over the course of several years in several separate individual differences investigations and experiments. Participants were split into three groups based on time of participation. Those who completed the laboratory tasks in Weeks 2 through 6 were grouped as Early participants; Weeks 7 through 11 as Middle participants; and Weeks 12 through 16 as Late participants (See Table 1 for the number of participants in each group)..

Procedure. Participants signed up for available participation times through an online scheduling system. Sessions lasted 2 hr.[1] Depending on the nature of the specific experiment, different groups of participants completed different combinations of tasks. The experiments included measures of working memory capacity (WMC), fluid intelligence (gF), crystallized intelligence (gC), long-term memory (LTM), and attention control (AC). For this and all subsequent experiments, we used tasks that have previously been used to measure the specific constructs of interest. In previous individual differences investigations, these tasks have shown both convergent and discriminant validity using confirmatory factor analyses (e.g., Unsworth, Fukuda, Awh, & Vogel, 2014). Scores on several measures of each ability were entered into a factor analysis and factor scores were saved, giving each participant a factor score for each cognitive ability. These factor scores were used in the statistical tests. Statistical tests on all individual measures are shown in the Appendix.

Table 1
Number of Participants in Experiment 1

Composite	Early	Middle	Late	Total
Working memory capacity	442	219	140	801
Fluid intelligence	385	136	111	632
Crystallized intelligence	293	61	60	414
Long-term memory	273	62	66	401
Attention control	138	147	63	348

Tasks.

Working memory capacity.

Operation span. The span tasks were used to measure working memory capacity because they require participants to both process and store information in working memory. In this task, participants solved a series of math operations while trying to remember a set of unrelated letters. Participants were required to solve a math operation, and after solving the operation, they were presented with a letter for 1 s. Immediately after the letter was presented the next operation was presented. At recall participants were asked to recall letters from the current set in the correct order by clicking on the appropriate letters. For all of the span measures, items were scored correct if the item was recalled correctly from the current list in the correct serial position. Participants were given practice on the operations and letter recall tasks only, as well as two practice lists of the complex, combined task. List length varied randomly from three to seven items. The score was total number of correctly recalled items.

Symmetry span. Participants recalled sequences of red squares within a matrix while performing a symmetry-judgment task. In the symmetry-judgment task, participants were shown an 8 × 8 matrix with some squares filled in black. Participants decided whether the design was symmetrical about its vertical axis. The pattern was symmetrical half of the time. Immediately after determining whether the pattern was symmetrical, participants were presented with a 4 × 4 matrix with one of the cells filled in red for 650 ms. At recall, participants recalled the sequence of red-square locations by clicking on the cells of an empty matrix. Participants were given practice on the symmetry-judgment and square recall task as well as two practice lists of the combined task. List length varied randomly from two to five items. We used the same scoring procedure as we used in the operation span task.

Reading span. While trying to remember an unrelated set of letters, participants were required to read a sentence and indicated whether or not it made sense. Half of the sentences made sense, while the other half did not. Nonsense sentences were created by changing one word in an otherwise normal sentence. After participants gave their response, they were presented with a letter for 1 s. At recall, participants were asked to recall letters from the current set in the correct order by clicking on the appropriate letters. Participants were given practice on the sentence judgment task and the letter recall task, as well as two practice lists of the combined task. List length varied randomly from three to seven items. We used the same scoring procedure as we used in the operation span and symmetry span tasks.

Long-term memory.

Delayed free recall. In this task participants were given six lists of 10 words each. All words were common nouns that were presented for 1 s each. After list presentation, participants engaged in a 16-s distractor task before recall. Participants saw 8 three-digit numbers appear on the screen for 2 s each and were required to

[1] Students are prevented from signing up for the same study more than once, even across academic terms. As each experiment used the same name on the online scheduling system, no participant contributed data more than once within or across experiments.

write the digits in ascending order. For example, if the participant saw "638" on the screen, they should write "368" on the sheet. After the distractor task participants saw "???" on the screen, which indicated that they should type as many words as they could remember from the current list in any order they wished. Participants had 45 s for recall. A participant's score was the total number of items recalled correctly. This task measures a person's ability to retrieve information from long-term memory after a delay.

Picture-source recognition. Participants were presented with a picture (30 total pictures) in one of four different quadrants on a screen for 1 s. Participants were explicitly instructed to pay attention to both the picture and the quadrant in which it was presented. At test, participants were presented with 30 old and 30 new pictures in the center of the screen. Participants were asked to indicate via key press whether the picture was new or old and, if the picture was old, the quadrant in which it was presented. Participants had 5 s to press the appropriate key to enter their response. A participant's score was proportion correct. This task measures a person's source memory.

Cued recall. In this task participants were given three lists of 10 word pairs each. All words were common nouns and the word pairs were presented vertically for 2 s each. Participants were told that the cue would always be the word on the top and the target would be at the bottom. After the presentation of the last word pair participants saw the cue word and "???" in place of the target word. Participants were instructed to type in the target word from the current list that matched the cue and then press "Enter" to indicate their response. The cues were randomly mixed so that the corresponding target words were not recalled in the same order as they were presented. Participants had 5 s to type in the corresponding word. The same procedure was done for all three lists. A participant's score was the proportion of items recalled correctly. This task measures a person's ability to retrieve cued information.

Attention control.

Antisaccade. The antisaccade task was used as a measure of attention control because participants must focus attention in order to avoid attentional capture by irrelevant stimuli. In this task (Kane et al., 2001) participants were instructed to stare at a fixation point which was onscreen for a variable amount of time (200 ms–2,200 ms). A flashing white "=" was then flashed either to the left or right of fixation (11.33° of visual angle) for 100 ms. This was followed by the target stimulus (a B, P, or R) onscreen for 100 ms. This was followed by masking stimuli (an H for 50 ms followed by an 8 which remained onscreen until a response was given). The participants' task was to identify the target letter by pressing a key for B, P, or R (the keys 1, 2, or 3) as quickly and accurately as possible. In the prosaccade condition the flashing cue (=) and the target appeared in the same location. In the antisaccade condition the target appeared in the opposite location as the flashing cue. Participants received, in order, 10 practice trials to learn the response mapping, 15 trials of the prosaccade condition, and 60 trials of the antisaccade condition. The dependent variable was proportion correct on the antisaccade trials.

Psychomotor vigilance task (PVT). The psychomotor vigilance task (PVT; Dinges & Powell, 1985) was used as the primary measure of sustained attention. Participants were presented with a row of zeros on screen and after a variable amount of time the zeros began to count up in 1-ms intervals from 0 ms. The participants' task was to press the spacebar as quickly as possible once the numbers started counting up. After pressing the space bar the response time was left on screen for 1 s to provide feedback to the participants. Interstimulus intervals were randomly distributed and ranged from 1 s to 10 s. The entire task lasted for 10 min for each individual (roughly 75 total trials). The dependent variable was the average reaction time (RT) for the slowest 20% of trials (Dinges & Powell, 1985).

Arrow flankers. The arrow flanker task was used as a measure of attention control because participants must ignore irrelevant stimuli and focus attention on the target. Participants were presented with a fixation point for 400 ms. This was followed by an arrow directly above the fixation point for 1,700 ms. The participants' task was to indicate the direction the arrow was pointing (pressing the F key for left pointing arrows and the J key for right pointing arrows) as quickly and accurately as possible. On 50 neutral trials the arrow was flanked by two horizontal lines on each side. On 50 congruent trials the arrow was flanked by two arrows pointing in the same direction as the target arrow on each side. Finally, on 50 incongruent trials the target was flanked by two arrows pointing in the opposite direction as the target arrow on each side. All trial types were randomly intermixed. The dependent variable was the RT difference between incongruent and congruent trials.

Fluid intelligence.

Raven advanced progressive matrices. The Raven is a measure of abstract reasoning and is commonly used in intelligence batteries (Raven et al., 1998). The test consists of 36 items presented in ascending order of difficulty (i.e., easiest to hardest). Each item consists of a display 3 × 3 matrices of geometric patterns with the bottom right pattern missing. The task for the participant is to select among eight alternatives the one that correctly completes the overall series of patterns. Participants received two practice items and were then given 10 min to complete the 18 odd-numbered items. A participant's score was the total number of correct solutions.

Verbal analogies. Analogies were used as a measure of verbal reasoning, which is another aspect of fluid intelligence. In this task participants read an incomplete analogy and were required to select the one word of five possible words that best completed the analogy. After one practice item, participants had 4 min to complete 18 test items. These items were originally selected from the Air Force Officer Qualifying Test (AFOQT; Berger, Gupta, & Skinner, 1990), and we used the same subset of items used in Kane et al. (2004). A participant's score was the total number of items solved correctly.

Number series. In this task participants saw a series of numbers and were required to determine what the next number in the series should be (Thurstone, 1962). That is, the series follows some unstated rule which participants are required to figure out in order to determine what the next number in the series should be. The ability to detect these patterns is a commonly noted element of fluid intelligence. Participants selected their answer from five possible numbers that were presented. Following five practice problems, participants had 3.5 min to complete 15 test times. A participant's score was the total number of items solved correctly.

Crystallized intelligence.

Synonym vocabulary. In this task participants were given 10 vocabulary words and were required to select the best synonym (of five possible choices) that best matched the target vocabulary word

(Hambrick, Salthouse, & Meinz, 1999). Participants were given 2 min to complete the 10 items. A participant's score was the total number of items solved correctly. This task, as well as antonym vocabulary, measures stored information about the meaning of words.

Antonym vocabulary. In this task participants were given 10 vocabulary words and were required to select the best antonym (of five possible choices) that best matched the target vocabulary word (Hambrick et al., 1999). Participants were given 2 min to complete the 10 items. A participant's score was the total number of items solved correctly.

General knowledge. In this task participants were given 24 general information questions and were required to select the best answer (of four possible choices) to the question (Hambrick et al., 1999). Topics included American politics, sports, music, literature, history, art, and economics. Participants were given 5 min to complete the 24 items. A participant's score was the total number of items solved correctly. This task measures stored information about a variety of topics.

Results and Discussion

For each factor, we ran an analysis of variance (ANOVA) with time of participation as a between-subjects factor. We also ran a *t* test to compare Early participants and Late participants to see if there are any differences between the earliest and latest participants, the "punctual" students and the "procrastinators."[2] Next, we ran a regression with week of participation as a continuous variable predicting each of the factors. Finally, we computed Bayes factor scores for each of the composite factors using the JASP software (www.jasp-stats.org; Love et al., 2015). A Bayes factor less than 1 can be interpreted as no evidence in favor an alternative model over the null hypothesis (Kass & Raftery, 1995). The results of all these tests are reported in Table 2 and are depicted graphically in Figure 1. As can be seen in the table, the *t*-statistics actually reveal that any possible effects are in the opposite direction of the hypothesized effect. That is, late participants actually outperform early participants on some tasks. Descriptive statistics and statistical tests on all individual measures are listed in the Appendix.

The results of the statistical tests in Table 2 offer several interesting findings, all of which counter the conventional wisdom. First, there is a slight difference between early and late participants in WMC such that late participants actually have slightly higher estimates. However, the Bayes factor did not indicate strong evidence in favor of the alternative hypothesis. Late participants actually score higher on measures of crystallized intelligence, which means they may actually be learning over the course of the academic term. The regression on fluid intelligence reached marginal significance ($p = .05$), but the relevant Bayes factor did not indicate evidence in favor of the alternative hypothesis. The squared standardized beta coefficient from the regression can also be used a measure of effect size. Using this analysis, the effect sizes were all right around 0. Also, the observed effects are actually in the opposite direction of a fatigue, task motivation, or "procrastination" hypothesis in which late participants score worse on measures of cognitive abilities for various reasons, and thus are inconsistent with such hypotheses.

Experiment 2

The second dataset comes from a series of experiments and individual differences investigations at the University of Oregon from a total of 1,594 participants. The university has four academic quarters which last for 10 weeks during the fall, winter, spring, and summer. The vast majority of data were collected during the fall, winter, and spring terms and all participants earned course credit for their participation. Participants completed various combinations of tasks measuring working memory capacity, fluid intelligence, and attention control. Additionally, thought probes were added to the attention control tasks, so there are also data on propensities to mind-wander and to experience external distraction during these tasks for 506 participants. Participants were again grouped by week of participation (see Table 3 for numbers of participants in each group).

Method

Participants. Data were collected over the course of several years. Participants signed up for available timeslots using an online scheduling system. Experiments typically took 2 hr to complete. Participants who completed the sessions in Weeks 2 through 4 of the academic term were grouped as Early participants; Weeks 5 through 7 as Middle participants; and Weeks 8 through 10 as Late participants.

Tasks.

Working memory capacity.

Operation span. See Experiment 1.

Symmetry span. See Experiment 1.

Reading span. See Experiment 1.

Fluid intelligence.

Raven advanced progressive matrices. See Experiment 1.

Letter sets. In this task participants saw five sets of four letters, and participants were required to induce a rule that applies to the composition and ordering of four of the five letter sets (Ekstrom, French, Harman, & Derman, 1976). Participants were then required to indicate the set that violates the rule. Following two examples participants had 5 min to complete 20 test items. A participant's score was the total number of items solved correctly.

Attention control.

Antisaccade. See Experiment 1. Thought probes followed 16% of antisaccade trials.

Psychomotor vigilance (PVT). See Experiment 1. Thought probes followed 20% of trials.

Sustained attention to response task (SART). Participants completed a version of a sustained attention to response task (SART) with semantic stimuli adapted from McVay and Kane (2009, 2012). The SART is a go/no-go task where subjects must respond quickly with a key press to all presented stimuli except infrequent (11%) target trials. In this version of the SART, word stimuli were presented in Courier New size 18 font for 300 ms followed by 900 ms mask. Most of the stimuli (nontargets) were members of one category (animals) and infrequent targets were

[2] We even further restricted an analysis comparing the earliest (Weeks 1 and 2) and the latest participants (Weeks 15 and 16). No significant differences arose for any of the other four factors (all *p*s > 0.06). Note: We could not always compare Week 2 and Week 16 because we did not have data for all measures from participants in these weeks. This test could also not be performed on the attention control measures because no participants completed the measures in Weeks 1 or 2.

Table 2
Results of Statistical Tests in Experiment 1

Composite	ANOVA (Time) *F*	*p*	BF	*t*-test (Early vs. Late) *t*	*p*	BF	Regression (Week) β	*t*	*p*	BF
WMC	1.91	.14	.053	−1.91	**.05**	.419	.04	1.06	.08	.137
gF	1.82	.16	.128	−1.48	.14	.314	.08	1.92	**.05**	.534
gC	2.78	.06	.509	−2.31	**.02**	**1.856**	.09	1.86	.06	.581
LTM	.36	.69	.053	−.02	.98	.150	.01	.28	.77	.115
AC	.51	.60	.055	.48	.63	.183	−.05	−.83	.41	.165

Note. ANOVA = analysis of variance; WMC = working memory capacity; gF = fluid intelligence; gC = crystallized intelligence; LTM = long-term memory; AC = attention control; BF = Bayes factor; β = standardized beta. Significant results are highlighted in bold.

members of a different category (foods). The SART had 470 trials, 50 of which were targets. The dependent variables were accuracy for targets and each individual's standard deviation of response time for go trials. Thought probes followed 60% of target trials. The task took approximately 10 min to complete.

Thought probes. During the attention control tasks, participants were periodically presented with thought probes asking them to classify their immediately preceding thoughts. The thought probes asked participants to press one of five keys to indicate what they were thinking just prior to the appearance of the probe. Specifically, participants saw:

Please characterize your current conscious experience.

1. I am totally focused on the current task

2. I am thinking about my performance on the task or how long it is taking

3. I am distraction by information present in the room (sights and sounds)

4. I am zoning out/my mind is wandering

5. Other

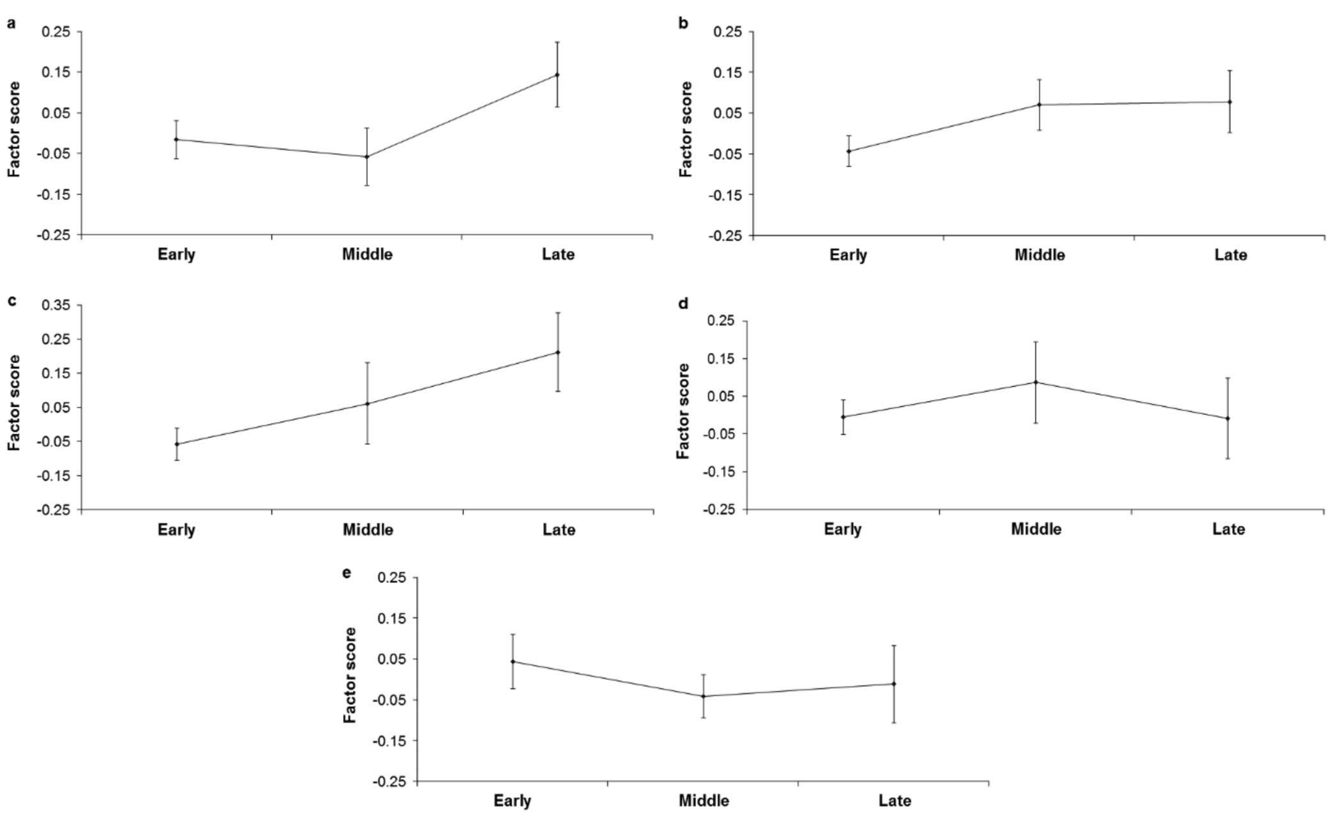

Figure 1. Results of Experiment 1 for (a) working memory capacity, (b) fluid intelligence, (c) crystallized intelligence, (d) long-term memory, and (e) attention control. Error bars represent ± one standard error of the mean.

Table 3
Number of Participants in Experiment 2

Composite	Early	Middle	Late	Total
Working memory capacity	500	538	556	1,594
Fluid intelligence	110	205	191	506
Attention control	110	205	190	505
Mind-wandering	110	205	191	506
External distraction	110	205	191	506

These thought probes were based on those used by Stawarczyk, Majerus, Maj, Van der Linden, and D'Argembeau (2011). During the instructions, participants were given specific instructions regarding the different categories. Similar to prior research, Responses 3 and 4 were considered task-unrelated thoughts, with Response 3 being classified as external distraction and Response 4 being classified as mind-wandering. Response 1 was considered on-task thought, while Response 2 was considered task-related interference. The present study only considers the reports of external distraction and mind-wandering in the analyses, as these two reports were of the most relevance to the current investigation.

Results and Discussion

The analyses are similar to those in Experiment 1. Results of the statistical tests are shown in Table 4 and are depicted graphically in Figure 2. ANOVAs with time of participation as a between-subjects factor revealed no significant differences between groups on any of the factors (WMC, gF, AC, external distraction, or mind-wandering). A follow-up *t* test comparing early and late participants revealed two significant effects. Early participants had slightly higher gF scores than late participants and slightly lower estimates of WMC.[3] Finally, treating week of participation as a continuous variable in a regression predicting these factors with week as the sole predictor revealed one counterintuitive effect. The later participants complete the working memory capacity measures, the higher their performance tended to be. Again, the observed effect sizes (squared standardized betas) were all lower than 0.01. In only one case (WMC) did the Bayes factor reach a considerable estimate. Most of the Bayes factors are below 1, which indicates evidence in favor the null hypothesis (no differences between groups). In order for a statistical result to be "substantial," the Bayes factor should be above 3.2 (Kass & Raftery, 1995). Most of the observed Bayes factors are nowhere near this threshold. No effects were observed for attention control abilities, propensities to mind-wander, or external distraction. Descriptive statistics and statistical tests on all measures are listed in the Appendix.

Experiment 2 revealed two effects that occurred in opposite directions: Late participants tended to have lower fluid intelligence scores and higher working memory capacity estimates. However, the lack of an observed effect in Experiment 1, the opposition of the two observed effects in Experiment 2, and the low Bayes factors of the observed effects in Experiment 2 do not clearly indicate that participants who complete laboratory tasks early in the term, in the middle of the term, and at the end of the term qualitatively differ in their cognitive abilities. Together, the results from Experiment 1 and Experiment 2 do not support the conventional wisdom that individuals who participate at different times during an academic term differ in any qualitative way in their cognitive abilities. Across a broad array of cognitive tasks, two different universities, two different academic schedules, and several experiments with large samples, there is little to no evidence that factors like fatigue, task motivation, or procrastination have any effect on late participants that may lead them to show poor cognitive performance.

Experiment 3

Experiments 1 and 2 investigated the possibility that individual differences in cognitive abilities vary systematically with time of participation. There was no strong evidence for this being the case. However, it is possible that individuals differ in their task motivation across the course of an academic term. At the beginning of an academic term, participants may be motivated to do well on the laboratory tasks. They want to do well in their courses so they have signed up early to complete their course requirements, and this motivation may carry over into the lab. Along the same vein, participants who enter the lab late in the term may feel unmotivated. They are simply participating to get a nuisance requirement out of the way, and thus they are unmotivated to perform well on laboratory tasks. There are certainly individual differences in task motivation, and these individual differences predict both performance and rates of mind-wandering during reading comprehension tasks (Robison & Unsworth, 2015; Unsworth & McMillan, 2013). Experiment 3 investigated this possibility. The current data came from experiments investigating the relationship between working memory capacity, reading comprehension, and mind-wandering. One of the dependent variables collected in both experiments was task motivation. So Experiment 3 utilizes this dataset to investigate the possibility that individuals differ in their motivation levels as a function of time of participation.

Method

Participants. Participants were 388 undergraduate students at the University of Oregon who participated for course credit. Participants were drawn from two experiments across the course of four academic quarters. Once again, participants were split into three groups in terms of time of participation: Weeks 2 to 4 as Early; Weeks 5 to 7 as Middle; and Weeks 8 to 10 as Late. See Table 5 for numbers of participants.

Tasks.

Working memory capacity tasks.

Operation span. See Experiment 1.

Symmetry span. See Experiment 1.

Reading span. See Experiment 1.

Text. Participants read an excerpt from The Challenge of Democracy (Janda, Berry, & Goldman, 2010), a popular introduc-

[3] We could not compare the most extreme groups (Weeks 2 and 10) for the fluid intelligence, attention control, external distraction, or mind-wandering measures because no participants completed these measures in Week 2. There was a significant effect for WMC ($p < .05$), but early participants actually scored lower, which is in the opposite direction of the typical time of participation effect. Note: This test compared 36 participants to 231 participants, but was still significant after an adjustment to the degrees of freedom (Leverne's test for inequality of variance).

Table 4
Results of Statistical Tests for Experiment 2

Composite	ANOVA (Time)			*t*-test (Early vs. Late)			Regression (Week)			
	F	*p*	BF	*t*	*p*	BF	β	*t*	*p*	BF
WMC	2.22	.10	.065	1.91	**.05**	.421	.07	2.72	**.01**	**2.193**
gF	2.37	.09	.218	−2.03	**.04**	.939	−.07	−1.76	.07	.451
AC	1.36	.25	.086	−.60	.54	.157	−.05	−1.15	.24	.189
MW	1.42	.24	.090	.11	.90	.130	−.002	−.04	.96	.099
ED	.75	.47	.047	1.19	.23	.261	.03	.69	.48	.125

Note. ANOVA = analysis of variance; WMC = working memory capacity; gF = fluid intelligence; AC = attention control; MW = reports of mind-wandering during attention control tasks; ED = reports of external distraction during attention control tasks; BF = Bayes factor; β = standardized beta. Significant effects are highlighted in bold.

tion to political science text. This text was chosen because it is an appropriate text for college students like the current sample. Each paragraph was presented in black font on a white background on the computer screen. When participants finished reading each paragraph, they proceeded to the next paragraph by pressing the space bar. There were 17 paragraphs in all, and participants were given as much time as needed to read the text.

Thought probes. Participants were given thought probes while they read the text in both experiments included in Experiment 2. But because there were subtle differences in the content of the probes across experiments, they are not included in the following analyses.

Reading comprehension test. Participants then completed a reading comprehension test based on the text. A total of 10 multiple choice questions comprised the test. Questions were taken directly from the online study guide accompanying the text. For each question, participants were asked about an aspect of the text (i.e., "What is the oldest objective of government?") and were required to respond by selecting one of five possible answers. Responses were scored by pressing the corresponding key. Participants were given as much time as needed to answer the questions. Scores were the percentage of questions answered correctly.

Questionnaire. After the reading comprehension test, participants were given a questionnaire to measure their interest in the

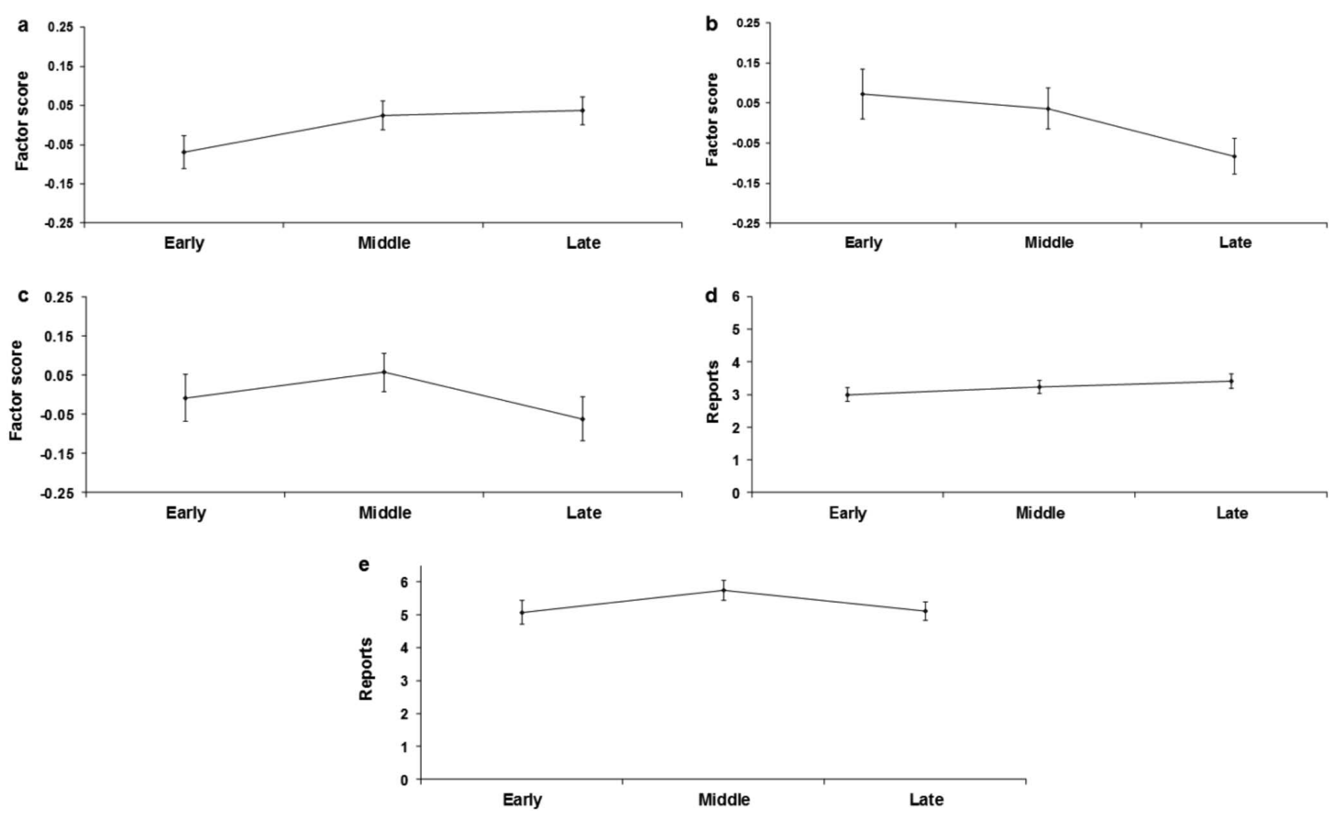

Figure 2. Results of Experiment 2 for (a) working memory capacity, (b) fluid intelligence, (c) attention control, (d) external distraction, and (e) mind-wandering. Error bars represent ± one standard error of the mean.

Table 5
Number of Participants in Experiment 3

Composite	Early	Middle	Late	Total
Working memory capacity	76	143	169	388
Reading comprehension	76	143	169	388
Motivation	76	143	169	388

topic, their experience with the topic, and their overall motivation to perform well. The two questions asking about task motivation were (a) "How motivated were you to do well on the task?" and (b) "How much did your motivation affect your performance on the task?" Participants responded by pressing the corresponding key on a scale of 1 (*not at all*) to 6 (*very much*). The primary dependent variable of interest in this experiment is task motivation, so the questions about topic interest and experience were not included in the analyses.

Results and Discussion

The dependent variables of interest were WMC, reading comprehension (RC), and task motivation. These variables were submitted to ANOVAs with time of participation as the sole between-subjects factor. A *t* test compared early participants with late participants,[4] and finally a regression treated week of participation as a continuous variable in predicting the three factors (WMC, RC, and motivation). The results are listed in Table 6 and shown graphically in Figure 3. No effects were observed for any of the three factors. Participants who completed the experiments early in the term, in the middle of the term, and at the end of the term performed roughly equally on the measures of working memory capacity and reading comprehension, and they also reported roughly equal levels of task motivation. Descriptive statistics and results for statistical tests on all measures are listed in the Appendix.

The results of Experiment 3 replicated the results found in Experiments 1 and 2 in that they offered no evidence that there are any differences in cognitive abilities between participants based on their time of participation. Experiment 3 extended those findings to the cognitive ability of reading comprehension. While there are substantial individual differences in working memory capacity and reading comprehension, neither of these individual differences covary systematically with time of participation.

These results do not lend support to the conventional wisdom that participants at the end of the term, the "procrastinators," are less motivated to complete laboratory tasks than the early participants, the "punctual." The conventional wisdom that the later a student completes a course requirement, in this case participation in a laboratory experiment, the less motivated they are, does not seem to hold up. One may question the validity of these motivation reports. However, Robison and Unsworth (2015) and Unsworth and McMillan (2013) showed that these motivation reports positively correlated with reading comprehension and negatively correlated with reports of mind-wandering. Although working memory capacity also positively correlated with reading comprehension and negatively correlated with mind-wandering, working memory capacity and motivation did *not* correlate. This suggests that working memory capacity and motivation offered independent contributions to mind-wandering and reading comprehension. The results of the current investigation seem to rule out the possibility that this effect interacts with time of participation. In the present study, participants seem to be roughly equally motivated across the course of the academic term.

Experiment 4

In addition to the cognitive measures used in Experiments 1, 2, and 3, the final experiment collected both measures of cognitive ability and personality. There is some evidence that individuals who sign up to participate at different times during the course of the academic term differ in terms of personality characteristics (Aviv et al., 2002; Harber et al., 2003; Stevens & Ash, 2001; Witt et al., 2011). This experiment collected data on working memory capacity, fluid intelligence, attention control, and personality across two academic terms at the University of Oregon.

Method

Participants. Data were collected from 160 participants, 137 of which completed all measures reported, over two consecutive academic terms from the undergraduate participant pool at the University of Oregon. All participants received course credit for participation. Participants were split into three groups in terms of time of participation: Weeks 2 to 4 as Early; Weeks 5 to 7 as Middle; and weeks 8 to 10 as Late.

Tasks.

Working memory capacity.

Operation span. The task delivered in Experiment 4 was a shortened form of the operation span task in Experiments 1, 2, and 3. Instead of receiving three lists of length 3, 4, 5, 6, and 7, participants only received two lists of each length. Otherwise the task was identical.

Symmetry span. The task was also a shortened form of the one reported in Experiments 1, 2, and 3. Instead of receiving three lists of length 2, 3, 4, and 5, participants received two lists of each length.

Reading span. The task was a shortened form of the one reported in Experiments 1, 2, and 3. Instead of receiving three lists of length 3, 4, 5, 6, and 7, participants only received two lists of each length.

Thought probes. The thought probes were identical to those used in experiment 2.

Attention control.

Antisaccade. See Experiment 2.

Psychomotor vigilance. See Experiment 2.

Stroop. Participants were presented with a color word (red, green, or blue) presented in one of three different font colors (red, green, or blue). The participants' task was to indicate the font color via key press (red = 1, green = 2, blue = 3). Participants were told to press the corresponding key as quickly and accurately as possible. Participants received 15 trials of response mapping practice and 6 trials of practice with the real task. Participants then received 100 real trials. Of these trials, 67% were congruent such that the word and the font color matched (i.e., red printed in red) and the

[4] Comparing only participants in Week 2 and Week 10 revealed no significant effects for any of the factors (all $ps > 0.37$).

Table 6
Results of Statistical Tests in Experiment 3

	ANOVA (time)			*t*-test (Early vs. Late)			Regression (Week)			
Composite	*F*	*p*	BF	*t*	*p*	BF	β	*t*	*p*	BF
WMC	1.48	.22	.120	−1.59	.11	.502	−.06	−1.18	.24	.224
RC	.28	.75	.039	.74	.45	.195	.07	1.44	.14	.308
Motivation	.12	.88	.034	−.45	.65	.165	−.03	−.61	.54	.135

Note. ANOVA = analysis of variance; WMC = working memory capacity; RC = reading comprehension; BF = Bayes factor; β = standardized beta.

other 33% were incongruent (i.e., red printed in green). The dependent variable was the RT difference between incongruent and congruent trials.

Fluid intelligence.

Raven advanced progressive matrices. See Experiment 1.

Letter sets. See Experiment 2.

Personality. Participants were given a shortened version of the Big Five Inventory questionnaire (John, Naumann, & Soto, 2008). The form has 44 items measuring the Big Five personality traits of extraversion, agreeableness, openness, neuroticism, and conscientiousness. Participants completed the form after completing all other measures.

Results and Discussion

We ran ANOVAs on working memory capacity, fluid intelligence, attention control, and the five factors of the Big Five Inventory (extraversion, agreeableness, conscientiousness, openness, and neuroticism), a follow-up *t* test comparing early participants and late participants,[5] and finally a regression treating week as a continuous variable. The results are shown in Table 7 and depicted graphically in Figure 4. Reliability estimates for the personality scales, descriptive statistics, and results for statistical tests for all individual measures are listed in the Appendix.

The results of Experiment 4 replicate Experiments 1, 2, and 3. There do not appear to be any differences in the cognitive abilities of individuals who participate early in the academic term, in the middle of the term, or at the end of the term. The results of the tests of the effect of time of participation on personality variables did not replicate prior research that individuals who participate at different times during the academic term qualitatively differ in their personality characteristics (Aviv et al., 2002; Stevens & Ash, 2001; Witt et al., 2011). From the Bayesian analyses, there is substantial evidence in favor of the null over the alternative hypothesis, and the largest observed effect size was 0.02. Despite this sample being smaller than some other studies (i.e., Witt et al., 2011), the sample size is large enough to have adequate power to detect reliable effects. Assuming a power of 0.80, the current sample size would be able to detect effects as small as 0.26 (see Table 8).

General Discussion

The conventional wisdom in psychological research on undergraduate participant pools is that toward the end of an academic term, participants suffer in their task performance for a number of reasons, which could potentially be differences in cognitive abilities, motivation levels, or personality characteristics. This introduces a confounding variable into the mix: the time of participation effect. However, utilizing a large database with over 2,900 participants across two universities, we found no convincing evidence that participants differ in any way in their cognitive abilities, motivation levels, or personality characteristics as a function of time of participation. Of course, even healthy young adults show substantial individual differences in cognitive abilities, as well as motivation levels and personality characteristics. However, none of these individual differences seem to covary with the time at which they participate in psychological experiments.

The first experiment investigated the time of participation effect on working memory capacity, attention control, long-term memory, fluid intelligence, and crystallized intelligence among an undergraduate population at the University of Georgia. The university has a 16-week academic semester with a 1-week break in the middle of each semester (Fall Break and Spring Break). None of the cognitive abilities measured showed any significant covariance with time of participation.

The second experiment largely replicated the results of Experiment 1 using data collected at the University of Oregon, which uses a 10-week quarter system. The analyses revealed one effect on WMC, which indicated that late participants actually have higher estimates. This finding is inconsistent with a fatigue or motivation account of time of participation effects, as estimates should be lower for late participants in these accounts. An effect on gF was also observed. However, this effect was in the opposite direction, the effect was small, and the Bayes factors for the statistical tests were not at traditional thresholds for substantiation. So together these observed effects do not offer any convincing evidence for systematic time of participation effects on cognitive abilities.

Experiment 3 investigated the possibility of differences in motivation levels across participants who complete laboratory tasks at various times in an academic term. Although motivation levels do correlate with task performance and reported rates of mind-wandering (Robison & Unsworth, 2015; Unsworth & McMillan, 2013), motivation levels did not covary with time of participation, nor did reading comprehension.

Finally, Experiment 4 investigated the possibility that participants differ in their personality traits as a function of time of

[5] We also compared the extreme groups of Week 2 and Week 10. Only agreeableness showed a significant effect, with early participants scoring as more agreeable than late participants ($p < .05$, all other $ps > 0.09$). Note: This test only compared eight participants with 14 participants.

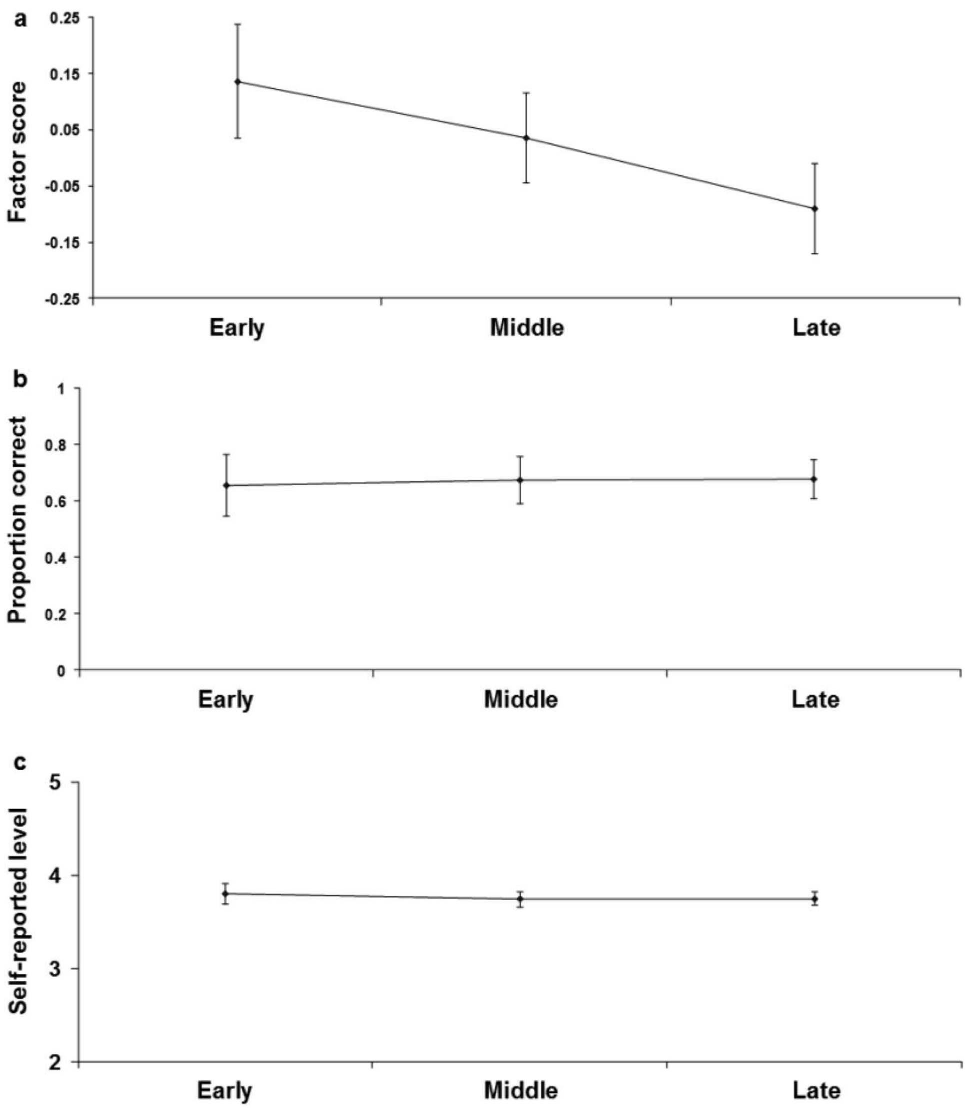

Figure 3. Results of Experiment 3 for (a) working memory capacity, (b) reading comprehension, and (c) motivation. Error bars represent ± one standard error of the mean.

participation. This effect has been observed several times (e.g., Aviv et al., 2002; Stevens & Ash, 2001; Witt et al., 2011) but was not replicated in the current sample. Admittedly, the current dataset (N = 137) was not as large as some studies that have observed this effect. However, if we committed a type-II error, the missed effect was small and perhaps unreliable.

The present study has several major advantages. First, with the exception of the personality dataset, the sample sizes are larger than any other studies that examined time of participation effects on cognitive performance. Second, the present study replicated findings across two universities with different academic schedules using similar and in most cases identical dependent measures. Neither a university population with a 16-week semester system nor a population with a 10-week quarter system showed any convincing evidence of time of participation effects on cognitive performance. Third, the present study had multiple measures of several different cognitive abilities including working memory capacity, long-term memory, attention control, fluid intelligence, crystallized intelligence, and propensity to mind-wander. This allowed us to look at cognitive performance from a domain-general level rather than task-specific level.

Arguing in favor of the null hypothesis is difficult. However, in this case the finding that participants' cognitive abilities, motivation levels, and personality characteristics do not covary systematically with time of participation is made possible because of the large sample sizes. In almost all analyses, the Bayes factors revealed strong evidence in favor of the null hypothesis over the alternative hypothesis. With the sample sizes in the present study, we are much more capable of making inferences about the population characteristics from which our samples are drawn, as the samples were drawn repeatedly from this population over the period of almost a decade.

Table 7
Number of Participants in Experiment 4

Composite	Early	Middle	Late	Total
Working memory capacity	54	52	54	160
Fluid intelligence	54	52	52	158
Attention control	50	45	48	143
Extraversion	47	45	45	137
Agreeableness	47	45	45	137
Conscientiousness	47	45	45	137
Openness	47	45	45	137
Neuroticism	47	45	45	137

Similarly, evidence in support of a null hypothesis is often uninteresting from a scientific perspective. However in this case, the results actually work to bolster the findings of most experiments and individual differences investigations that utilize undergraduate participant pools for sampling. Because time of participation does not seem to be a significant confound, experimenters need not be overly concerned about this often unmeasured and unaccounted for source of variance. Rather, experimenters can be confident in carrying out experiments throughout an academic term, which grants them greater flexibility in data collection.

We would be remiss to not acknowledge several of the limitations of the findings. First, the data is collected entirely from undergraduate populations at two large state universities in the U.S. Therefore, there could be a restriction of the range that led to an inability to detect time of participation effects. Typically, to extend the range cognitive researchers will also recruit community members that are not undergraduate students at their respective university. But because the time of participation effect is specifically tied to the academic term, we felt comfortable including only students in our design. We are also confident in the diversity of our samples from these two universities, as even at rather selective

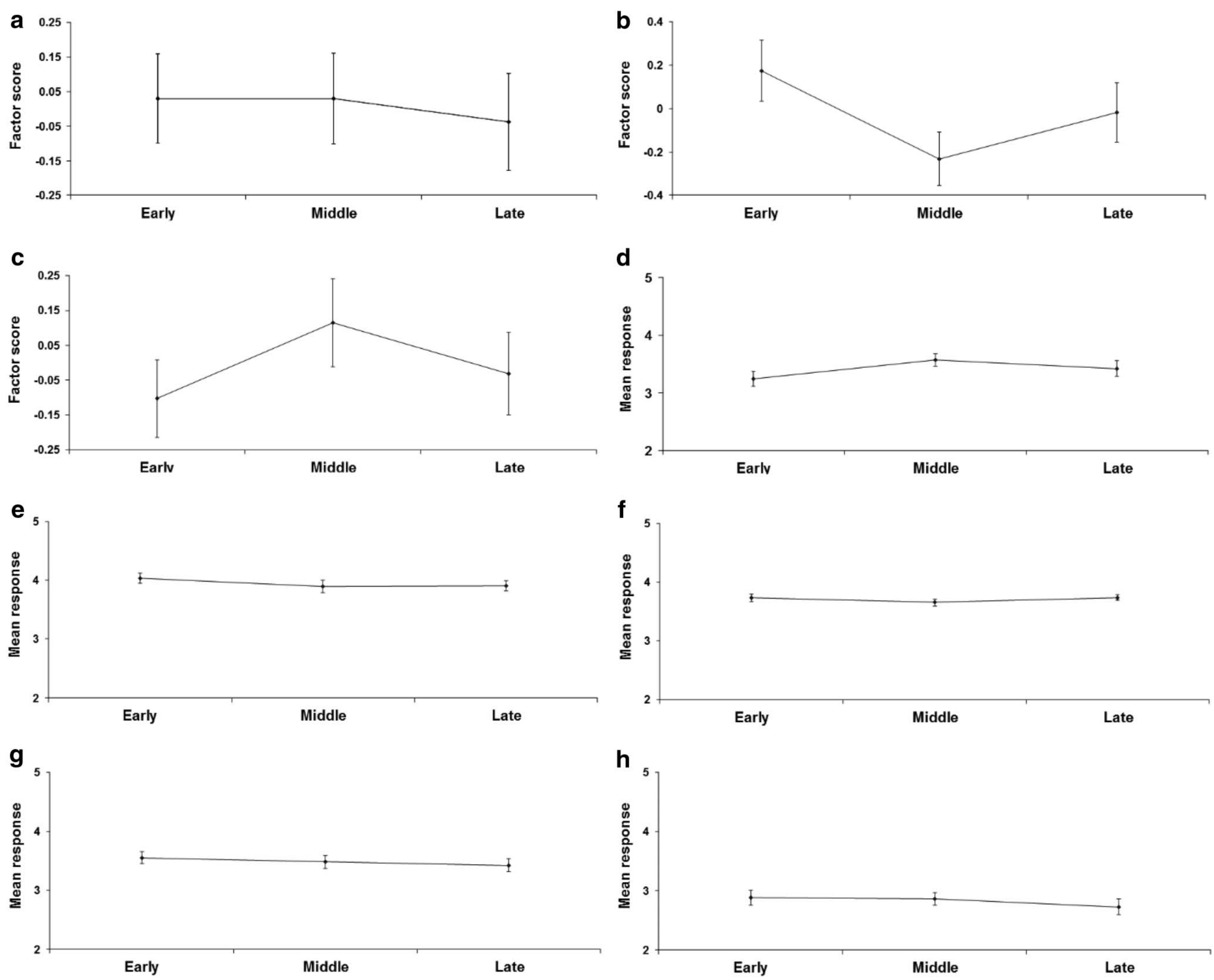

Figure 4. Results of Experiment 4 for (a) working memory capacity, (b) fluid intelligence, (c) attention control, (d) extraversion, (e) agreeableness, (f) conscientiousness, (g) openness, and (h) neuroticism. Error bars represent ± one standard error of the mean.

Table 8
Results of Statistical Tests in Experiment 4

Composite	ANOVA (Time)			*t*-test (Early vs. Late)			Regression (Week)			
	F	*p*	BF	*t*	*p*	BF	β	*t*	*p*	BF
WMC	.09	.91	.067	−.36	.71	.216	−.06	−.79	.43	.229
gF	2.33	.10	.460	−.98	.32	.316	−.10	−1.31	.18	.384
AC	.86	.42	.142	.44	.65	.232	.01	.15	.88	.183
Extra	1.76	.17	.313	.96	.34	.328	.04	.51	.61	.208
Agree	.70	.49	.128	−1.06	.29	.360	−.08	−.88	.36	.264
Cons	.61	.54	.119	.03	.96	.219	.02	.29	.77	.192
Open	.63	.53	.122	−1.09	.27	.369	−.16	−1.89	.06	.896
Neur	.63	.53	.121	−1.06	.28	.362	−.06	−.71	.48	.232

Note. ANOVA = analysis of variance; WMC = working memory capacity; gF = fluid intelligence; AC = attention control; Extra = extraversion; Agree = agreeableness; Cons = conscientiousness; Open = openness; Neur = neuroticism.

universities participants show considerable variability in cognitive abilities and personality characteristics. Second, we acknowledge that we only included Big Five measures of personality. We recognize that there are different conceptualizations of personality profiles that we did not include, and there are potential differences in these conceptualizations that we could not observe. Third, we cannot address the notion of regulatory fit or focus, because we did not measure promotion- and prevention-focused motivational states. Rather, our questions were intended to measure task-level motivation, which consistently correlates with task performance and not with cognitive abilities (Robison & Unsworth, 2015; Unsworth & McMillan, 2013). Finally, we recognize that there is no strong theory for why time of participation effects may occur and thus it is difficult to refute such a theory. However there are several potential explanations for how these effects may arise: (a) systematic differences in cognitive abilities such that earlier participants have better abilities, (b) motivational differences in which the late participants have lower motivation to perform well on laboratory tasks, and (c) early and late participants have different personality profiles. We provide strong evidence that cognitive abilities, task motivation, and Big Five personality traits do not consistently differ from the beginning to the end of the semester among undergraduate participants. Therefore, researchers who use undergraduate participant pools need not be overly concerned with such a confound and can collect data across the full academic term.

Researchers should always be careful to ensure random assignment to experimental conditions, proper counterbalancing, and the consideration of potential confounding variables. Some tasks may indeed have idiosyncrasies that lend themselves to time of participation effects. As can be seen in Table A1 of the Appendix, in some experiments, some tasks do indeed show time of participation effects. However, none of the task-specific effects replicate from one experiment to another. This underscores the importance of using several tasks to measure a latent construct, especially in individual differences research. It also underscores the importance of replication when these effects do indeed occur. Finally, it shows that in some cases a sample may indeed show a time of participation effect when there is only one dependent variable in an experiment. If researchers have reason to believe this effect is present in their sample, statistical steps (such as ANCOVA, mediation, and partial correlations) can be used to partial out variance due to this effect. If in the specific task it seems as though time of participation may be affecting the dependent variable of interest, researchers should measure this effect and account for it in their analyses. One additional caveat of the present study is that all participants received course credit for participation. As Nicholls et al. (2015) observed, credit participants and paid participants may differ in their intrinsic and extrinsic motivation levels, which could interact with the dependent variable. So this study could not address the potential confound of various forms of compensation. If some participants in an experiment receive course credit and others are paid, experimenters should be careful that these different compensation systems do not inadvertently affect the results.

Conclusion

Utilizing data from over 2,900 participants at two universities with two academic schedules across several years, the current results do not suggest that there is any reason to be overly concerned with systematic differences in cognitive abilities, motivation levels, or personality characteristics as a function of time of participation.

References

Amabile, T. M., Hill, K. G., Hennessey, B. A., & Tighe, E. M. (1994). The Work preference inventory: Assessing intrinsic and extrinsic motivational orientations. *Journal of Personality and Social Psychology, 66,* 950–967. http://dx.doi.org/10.1037/0022-3514.66.5.950

Aviv, A. L., Zelenski, J. M., Rallo, L., & Larsen, R. J. (2002). Who comes when: Personality differences in early and later participation in a university subject pool. *Personality and Individual Differences, 33,* 487–496. http://dx.doi.org/10.1016/S0191-8869(01)00199-4

Bender, T. (2007). Time of participation effect and grade orientation. *Personality and Individual Differences, 43,* 1175–1183.

Berger, F. R., Gupta, W. B., & Skinner, J. (1990). *Air Force Officer Qualifying Test (AFOQT) form P: Test manual (AFHRL-TR-89*–56). Brooks Air Force Base, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Blatt, S. J., & Quinlan, P. (1967). Punctual and procrastinating students: A study of temporal parameters. *Journal of Consulting Psychology, 31,* 169–174. http://dx.doi.org/10.1037/h0024413

Costa, P. T., & McRae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

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. http://dx.doi.org/10.3758/BF03200977

Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). Kit of factor-referenced cognitive tests (Rev. Ed.). Princeton, N. J.: Educational Testing Service.

Grimm, L. R., Markman, A. B., & Maddox, W. T. (2012). End-of-semester syndrome: How situational regulatory fit affects test performance over an academic semester. *Basic and Applied Social Psychology, 34,* 376–385. http://dx.doi.org/10.1080/01973533.2012.693427

Hambrick, D. Z., Salthouse, T. A., & Meinz, E. J. (1999). Predictors of crossword puzzle proficiency and moderators of age-cognition relations. *Journal of Experimental Psychology: General, 128,* 131–164. http://dx.doi.org/10.1037/0096-3445.128.2.131

Harber, K. D., Zimbardo, P. G., & Boyd, J. N. (2003). Participant self-selection biases as a function of individual differences in time perspective. *Basic and Applied Social Psychology, 25,* 255–264. http://dx.doi.org/10.1207/S15324834BASP2503_08

Janda, K., Berry, J. M., & Goldman, J. (2010). *The challenge of democracy*. Boston, MA: Houghton Mifflin Company.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. In O. P. John, R. W. Robbins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research*. New York, NY: Guilford Press.

Johnson, J. A. (2000). *Development of a short form of the IPIP-NEO: A report to HCW Consulting*. Unpublished manuscript, Department of Psychology, University of Pennsylvania, DuBois, PA.

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

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133,* 189–217. http://dx.doi.org/10.1037/0096-3445.133.2.189

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90,* 773–795. http://dx.doi.org/10.1080/01621459.1995.10476572

Kleinsorge, T., Diestel, S., Scheil, J., & Niven, K. (2014). Burnout and the fine-tuning of cognitive resources. *Applied Cognitive Psychology, 28,* 274–278. http://dx.doi.org/10.1002/acp.2999

Langston, W., Ohnesorge, C., Kruley, P., & Haase, S. J. (1994). Changes in subject performance during the semester: An empirical investigation. *Psychonomic Bulletin & Review, 1,* 258–263. http://dx.doi.org/10.3758/BF03200778

Love, J., Selker, R., Marsman, M., Jamil, T., Verhagen, A. J., Ly, A., . . . Wagenmakers, E.-J. (2015). JASP (Version 0.6.6) [Computer software].

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 196–204. http://dx.doi.org/10.1037/a0014104

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General, 141,* 302–320. http://dx.doi.org/10.1037/a0025250

Nicholls, M. E. R., Loveless, K. M., Thomas, N. A., Loetscher, T., & Churches, O. (2015). Some participants may be better than others: Sustained attention and motivation are higher early in semester. *The Quarterly Journal of Experimental Psychology, 68,* 10–18. http://dx.doi.org/10.1080/17470218.2014.925481

Page, M. M., & Lumia, A. R. (1968). Cooperation with demand characteristics and the bimodal distribution of verbal conditioning data. *Psychonomic Science, 12,* 243–244.

Raven, J. C., Raven, J. E., & Court, J. H. (1998). *Progressive matrices*. Oxford, UK: Oxford Psychologists Press.

Richter, D. O., Wilson, S. D., Milner, M., & Senter, R. J. (1981). Some differences among students volunteering as research subjects. *Bulletin of the Psychonomic Society, 17,* 261–263.

Robison, M. K., & Unsworth, N. (2015). Working memory capacity offers resistance to mind-wandering and external distraction in a context-dependent manner. *Journal of Applied Cognitive Psychology, 29,* 680–690.

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. http://dx.doi.org/10.1016/j.actpsy.2011.01.002

Stevens, C. D., & Ash, R. A. (2001). The conscientiousness of students in subject pools: Implications for "laboratory" research. *Journal of Research in Personality, 35,* 91–97. http://dx.doi.org/10.1006/jrpe.2000.2310

Thurstone, L. L. (1962). *Primary mental abilities*. Chicago, IL: Science Research Associates.

Underwood, B. J., Schwenn, E., & Keppel, G. (1964). Verbal learning as related to point of time in the school term. *Journal of Verbal Learning and Verbal Behavior, 3,* 473–480. http://dx.doi.org/10.1016/S0022-5371(64)80045-1

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology, 71,* 1–26. http://dx.doi.org/10.1016/j.cogpsych.2014.01.003

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory & Cognition, 39,* 832–842. http://dx.doi.org/10.1037/a0029669

Wang, A. Y., & Jentsch, F. G. (1998). Point-of-time effects across the semester: Is there a sampling bias? *The Journal of Psychology,* 211–219.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York, NY: Psychological Corporation.

Witt, E. A., Donnellan, M. B., & Orlando, M. J. (2011). Timing and selection effects within a psychology subject pool: Personality and sex matter. *Personality and Individual Differences, 50,* 355–359. http://dx.doi.org/10.1016/j.paid.2010.10.019

Appendix
Statistical Tests for all Measures

Table A1
Statistical Tests for all Individual Measures

	ANOVA (Time)			*t*-test (Early vs. Late)			Regression (Week)			
Measure	*F*	*p*	BF	*t*	*p*	BF	β	*t*	*p*	BF
					Experiment 1					
OSpan	.62	.53	.03	−.21	.83	.21	−.01	−.23	.82	.081
RSpan	.66	.51	.03	1.76	.08	1.18	.03	.76	.45	.104
SymSpan	4.20	**.02**	.87	1.98	**.05**	.82	.08	2.29	**.03**	.818
DFR	.16	.85	.03	1.35	.18	.47	.02	.48	.63	.102
PicSourceRec	3.12	**.04**	.32	2.01	**<.01**	**10.99**	.08	2.35	**.02**	**1.204**
Cued recall	5.60	**.01**	**4.71**	−2.31	.03	**10.391**	−.14	−3.30	**.01**	**17.712**
Antisaccade	.35	.70	.05	.12	.91	.164	.05	.89	.37	.172
Flanker	1.08	.33	.09	−1.40	.16	.408	−.07	−1.31	.19	.266
PVT	.46	.63	.05	.44	.65	.180	.03	.49	.63	.132
RAPM	3.60	**.03**	.48	.28	.78	.21	.02	−.02	.47	.104
Number series	2.43	.09	.23	.77	.44	.27	**.08**	**2.13**	**.03**	.811
Analogies	1.17	.31	.07	2.25	**.02**	**1.93**	.06	1.53	.13	.277
Syn. vocab	3.57	**.03**	**1.06**	1.35	.18	.48	.11	2.15	**.03**	**1.002**
Ant. vocab	1.71	.18	.18	.45	.65	.24	.06	1.18	.24	.212
Gen. knowledge	.78	.45	.08	.82	.41	.29	.05	.92	.36	.163
					Experiment 2					
OSpan	1.24	.29	.02	1.30	.19	.16	.06	2.23	**.03**	.659
RSpan	3.03	**.05**	.15	2.34	**.02**	1.08	.07	2.77	**.01**	**2.458**
SymSpan	.19	.83	.01	.42	.68	.08	.02	.80	.43	.077
Antisaccade	1.36	.26	.08	1.58	.12	.43	.06	1.38	.17	.248
PVT	4.64	**.01**	**1.84**	2.01	**.05**	.89	.13	.13	**.01**	**8.423**
SART	1.14	.32	.07	.92	.36	.20	.06	1.41	.16	.260
RAPM	1.35	.26	.08	−1.47	.14	.37	−.05	−1.02	.31	.164
Letter sets	1.75	.18	.12	−1.72	.09	.53	−.08	−1.83	.07	.498
					Experiment 3					
OSpan	.98	.38	.08	−1.19	.23	.29	−.04	−.86	.39	.161
RSpan	2.45	.09	.28	−2.18	**.03**	1.38	−.08	−1.66	.10	.421
SymSpan	.24	.79	.04	−.54	.59	.17	−.02	−.34	.73	.119
Read Comp.	.28	.76	.04	.74	.46	.20	.07	1.45	.54	.308
Motivation	.13	.88	.03	−.45	.65	.17	−.03	−.61	.54	.135
					Experiment 4					
OSpanR	.04	.96	.07	−.11	.91	.21	−.03	−.32	.75	.180
RSpanR	.27	.76	.08	−.02	.98	.20	−.03	−.33	.74	.180
SymSpanR	.49	.61	.10	−.77	.45	.26	−.10	−1.23	.22	.345
Antisaccade	.15	.86	.07	−.35	.72	.22	−.01	−.08	.93	.173
Stroop	.90	.41	.14	−1.25	.22	.41	−.12	−1.44	.15	.447
PVT	1.04	.36	.16	−.41	.69	.23	−.03	−.31	.76	.185
RAPM	3.42	**.04**	**1.15**	−1.55	.13	.59	−.14	−1.71	.09	.650
Letter sets	.50	.61	.10	−.16	.88	.21	−.04	−.47	.64	.191
Extraversion	1.77	.18	.31	.96	.34	.33	.04	.51	.61	.208
Agreeableness	.70	.50	.13	−1.06	.29	.36	−.08	−.89	.38	.264
Conscientiousness	.62	.54	.12	.04	.97	.22	.03	.29	.77	.192
Openness	.64	.53	.12	−1.09	.28	.37	−.16	−1.87	.06	.896
Neuroticism	.63	.53	.12	−1.07	.29	.36	−.06	−.71	.48	.232

Note. ANOVA = analysis of variance; OSpan = operation span; RSpan = reading span; SymSpan = symmetry span; DFR = delayed free recall; PicSourceRec = picture source recognition; PVT = psychomotor vigilance task; RAPM = Raven advanced progressive matrices; Syn. vocab. = synonym vocabulary; Ant. vocab. = antonym vocabulary; Gen. knowledge = general knowledge; Read Comp = reading comprehension; OSpanR = reduced operation span; RSpanR = reduced reading span; SymSpanR = reduced symmetry span. Significant results are highlighted in bold.

(*Appendix continues*)

Table A2
Means and Standard Deviations for all Measures

Factors	Measure	Early	Middle	Late
		Experiment 1		
WMC	Operation span	60.40 (11.23)	59.59 (12.83)	60.95 (11.90)
	Symmetry span	29.43 (8.29)	29.21 (7.60)	31.44 (6.30)
	Reading span	59.97 (13.63)	56.82 (13.76)	58.37 (13.69)
gF	RAPM	10.36 (2.43)	10.78 (2.72)	10.09 (2.61)
	Number series	9.35 (2.53)	9.84 (2.48)	9.76 (2.63)
	Analogies	11.20 (2.98)	11.21 (2.76)	11.68 (3.24)
gC	Synonym vocab.	4.05 (1.97)	4.49 (2.49)	4.23 (2.14)
	Antonym vocab.	4.29 (1.98)	4.23 (1.24)	4.80 (2.20)
	General knowledge	15.54 (3.69)	15.02 (4.16)	15.07 (3.35)
LTM	Delayed free recall	32.43 (7.91)	32.44 (9.00)	32.93 (9.04)
	Picture source rec.	.73 (.18)	.74 (.17)	.78 (.16)
	Cued recall	.40 (.25)	.33 (.23)	.34 (.27)
AC	Antisaccade	.52 (.14)	.53 (.14)	.52 (.14)
	Flanker	128.12 (74.11)	121.10 (65.03)	113.51 (54.51)
	PVT	600.02 (441.78)	574.08 (293.80)	632.92 (574.78)
		Experiment 2		
WMC	Operation span	56.38 (13.46)	57.55 (10.91)	56.48 (12.61)
	Symmetry span	21.60 (9.17)	23.49 (9.98)	23.30 (9.54)
	Reading span	37.08 (18.02)	40.89 (19.29)	41.42 (18.09
AC	Antisaccade	.46 (.12)	.48 (.12)	.48 (.11)
	PVT	576.63 (163.58)	572.59 (176.18)	634.71 (276.83)
	SART	150.34 (42.47)	148.79 (43.15)	155.46 (48.72)
gF	RAPM	8.58 (2.91)	8.48 (3.13)	8.07 (2.82)
	Letter sets	10.04 (3.01)	9.89 (2.98)	9.46 (2.69)
		Experiment 3		
WMC	Operation span	58.51 (9.97)	58.13 (11.88)	56.62 (12.12)
	Symmetry span	29.58 (7.08)	29.52 (6.88)	29.05 (7.16)
	Reading span	57.26 (11.97)	54.54 (13.19)	53.13 (14.44)
—	Reading comp.	.65 (.18)	.67 (.20)	.67 (.20)
—	Motivation	3.80 (.96)	3.73 (1.01)	3.74 (.89)
		Experiment 4		
WMC	OSpan reduced	36.90 (9.12)	37.13 (7.61)	36.36 (9.75)
	SymSpan reduced	19.87 (5.02)	18.94 (4.95)	19.60 (6.96)
	RSpan reduced	36.21 (9.07)	37.40 (9.02)	36.58 (10.23)
AC	Antisaccade	.62 (.15)	.64 (.18)	.63 (.16)
	PVT	464.43 (88.70)	440.59 (83.69)	457.66 (79.29)
	Stroop	173.82 (90.88)	172.32 (99.93)	152.03 (88.75)
gF	RAPM	9.67 (2.98)	8.21 (2.81)	8.79 (2.92)
	Letter sets	10.14 (3.09)	9.59 (3.12)	10.05 (2.91)
—	Extraversion	3.24 (.89)	3.57 (.73)	3.42 (.90)
—	Agreeableness	4.03 (.58)	3.89 (.70)	3.90 (.58)
—	Conscientiousness	3.72 (.43)	3.65 (.39)	3.74 (.34)
—	Openness	3.55 (.68)	3.48 (.50)	2.42 (.41)
—	Neuroticism	2.88 (.68)	2.86 (.77)	2.72 (.75)

Note. Standard deviations are in parentheses. WMC = working memory capacity; LTM = long-term memory; AC = attention control; gF = fluid intelligence; gC = crystallized intelligence; Picture source rec. = picture source recognition; PVT = psychomotor vigilance task; RAPM = Raven advanced progressive matrices; Synonym vocab. = synonym vocabulary; Antonym vocab. = antonym vocabulary; SART = sustained attention to response task; Reading comp. = reading comprehension; OSpan reduced = reduced version of operation span; SymSpan reduced = reduced version of symmetry span; RSpan reduced = reduced version of reading span.

(*Appendix continues*)

Table A3
Reliability Coefficients for Personality Measurements

Factor	No. of items	Cronbach's α
Extraversion	8	.88
Agreeableness	9	.79
Conscientiousness	9	.69
Openness	8	.80
Neuroticism	10	.69

Received June 5, 2015
Revision received September 9, 2015
Accepted September 30, 2015 ■