IN PRESS:
JOURNAL OF EXPERIMENTAL PSYCHOLOGY: GENERAL

A Combined Experimental and Individual-Differences Investigation into Mind Wandering During a Video Lecture

Michael J. Kane and Bridget A. Smeekens

University of North Carolina at Greensboro

Claudia C. von Bastian, John H. Lurquin, Nicholas P. Carruth, and Akira Miyake

University of Colorado Boulder

Author Notes

Michael J. Kane and Bridget A. Smeekens, Department of Psychology, University of North Carolina at Greensboro; Claudia C. von Bastian, John H. Lurquin, Nicholas P. Carruth, and Akira Miyake, Department of Psychology and Neuroscience, University of Colorado Boulder. Claudia C. von Bastian is now at Bournemouth University.

For assistance in data collection, we thank Breann Donnolly, Christopher Farmelo, Caitlin Graham, Jonathan Parrish, Joshua Perkins, Natalie Phillips, Laura Tran, Juan Ventura, and Adam Winter. For assistance in study management, we thank Caitlin Graham. For assistance in scoring posttest Parts 2 and 3, we thank Natalie Phillips. For valuable input during the design of the study, we thank Andrew Conway, Judith Harackiewicz, Lisa Linnenbrink-Garcia, Marsha Lovett, and Paul Silvia.

This research was supported by award numbers DRL1252333 (to Michael Kane) and DRL1252385 (to Akira Miyake) from the National Science Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation.

The stimulus materials (e.g., the lecture videos, the pretest/posttest stats questions) as well as annotated data files from this study are available on the following URL: [URL info here].

Correspondence should be addressed to either Michael J. Kane, Department of Psychology, P.O. Box 26170, University of North Carolina at Greensboro, Greensboro, NC 27402-6170; e-mail: mjkane@uncg.edu, or Akira Miyake, Department of Psychology and Neuroscience, University of Colorado Boulder, 345 UCB, Boulder, CO 80309-0345; email: akira.miyake@colorado.edu.

Abstract

A combined experimental–correlational study with a diverse sample (*N* = 182) from two research sites tested a set of five *a priori* hypotheses about mind wandering and learning, using a realistic video lecture on introductory statistics. Specifically, the study examined whether students' vulnerability to mind wandering during the lecture would predict learning from, and situational interest in, the video, and also whether longhand note-taking would help reduce mind wandering, at least for some students. Half the subjects took notes during the video, and all were subsequently tested on lecture content without notes. Regression and mediation analyses indicated that: (a) several individual-differences variables (e.g., pretest score, prior math interest, classroom media multitasking habits) uniquely predicted in-lecture mind wandering frequency; (b) although the note-taking manipulation did not reduce mind wandering at the group level, note-taking still reduced mind wandering for some individuals (i.e., those with lower prior knowledge and those who took notes of high quality and quantity); (c) mind wandering uniquely predicted both learning (posttest) and situational interest outcomes above and beyond all other individual-differences variables; (d) moreover, mind wandering significantly mediated the effects of several individual differences; and, finally, (e) not all types of mind wandering were problematic—in fact, off-task reflections about lecture-related topics positively predicted learning. These results, which were generally robust across the two sites, suggest that educationally focused cognitive research may benefit from considering attentional processes during learning as well as cognitive and noncognitive individual differences that affect attention and learning.

Keywords: mind wandering, note-taking, learning, education, situational interest, media multitasking

A Combined Experimental and Individual-Differences Investigation into Mind Wandering During a Video Lecture

How can teachers and students optimize learning? Cognitive psychologists are increasingly applying laboratory findings to this educational problem. Principles from the memory literature, such as the benefits of spaced practice, testing, and metacognitive self-evaluations, have been most successfully employed and broadly disseminated (for reviews, see Benassi, Overson, & Hakala, 2014; Brown, Roediger, & McDaniel, 2014; Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). We suggest that this education-focused research should also embrace two additional domains: (a) individual differences in cognitive abilities and domain experience (e.g., interest and knowledge) and (b) the functions—and dysfunctions—of attention in learning contexts.

Stable cognitive-ability differences, such as those in working memory capacity (WMC), predict many indices of academic achievement (e.g., Alloway & Alloway, 2010; Cowan et al., 2005). Educational psychology has further identified personality and experience constructs linked to scholastic success, such as conscientiousness, self-efficacy, and mastery goals (e.g., Pintrich, 2003; Richardson, Abraham, & Bond, 2012; Robbins et al., 2004). Thus, learning-science research motivated by cognitive psychology should explore individual-by-treatment interactions (Cronbach, 1957; Snow, 1989), whereby the adoption and effectiveness of learning and instructional strategies vary across students who differ along theoretically relevant dimensions.

Educationally relevant research should also pay more attention to attention. Although classrooms are designed to limit external distractions, the ubiquity of internet-connected devices can thwart these designs and disrupt student focus (e.g., Hembrooke & Gay, 2003; Ravizza, Uitvlugt, & Fenn, 2017). Of course, students can also be distracted even without alluring stimuli. *Mind wandering*, which refers to daydreaming, worrying, or other forms of off-task (or context-independent) thinking, is increasingly appreciated as a frequent and important cognitive activity that confers both costs and benefits (e.g., Immordino-Yang, Christodoulou, & Singh, 2012; Kane & McVay, 2012; Smallwood & Schooler, 2015).

The present study combined experimental and individual-differences approaches to identify predictors and possible consequences of mind wandering during learning, using a realistic video lecture. Within a laboratory context, we assessed a wide range of individual differences variables (e.g., WMC, background knowledge and interest, epistemic beliefs) to test five *a priori* hypotheses (**H1**–**H5**) about mind wandering and learning derived from the existing literature. Specifically, we tested: (a) which relevant individual-differences variables uniquely predicted students' mind wandering rates during the lecture (**H1**); (b) whether these individual differences interact with an experimental note-taking treatment to affect mind wandering (**H2**); (c) whether students' mind wandering rates predicted two outcome variables—posttest scores and situational interest—above and beyond other variables (**H3**); (d) to what extent mind wandering rates mediated the associations between individual-differences factors (and any individual-by-treatment interactions) and the outcome variables (**H4**); and, finally, (e) whether some forms of mind wandering might actually be beneficial to learning (**H5**).

Our joint testing of these five hypotheses makes the current study unique, especially in light of the limited existing research on mind wandering that has rigorously examined individual differences in educationally relevant settings. Before elaborating the five hypotheses, however, we briefly review relevant research on mind wandering and on note-taking, while highlighting the key findings and limitations that motivated the current study.

Mind Wandering in Educationally Relevant Contexts

Laboratory and classroom studies typically measure mind wandering by periodically interrupting subjects' ongoing activities to signal them to report on their immediately preceding thoughts. These thought probes require subjects to categorize their thoughts as on- versus off-task, or choose among a few categories to indicate what they were just thinking about. Thought probe responses are used to assess momentary correlates of *task-unrelated thoughts* (TUTs) and individual differences in TUT rates. Most of the educationally relevant research suggests that more frequent mind wandering is associated with worse comprehension and less learning.

Mind Wandering While Reading

In studies of mind wandering while reading, subjects read texts and answer questions about them. Thought probes appear unpredictably during the text to assess momentary correlates (e.g., Feng, D'Mello, & Graesser, 2013; Smallwood, McSpadden, & Schooler, 2008) and individual-differences correlates (e.g., Grodsky & Giambra, 1990; Varao Sousa, Carriere, & Smilek, 2013) of TUTs with comprehension. For example, regarding momentary correlates, when probes are followed by a test for the just-presented material, mind-wandering reports coincide with poorer recall of the preceding text than do on-task reports (Schooler, Reichle, & Halpern, 2004). As an example of individual-differences correlates, people who report more TUTs at probes during reading also tend to recall less of what they read than do those who report fewer TUTs (Schooler et al., 2004). Another important conclusion from individual-differences studies is that TUT rates are reliable: Students who mind-wander more during one reading task also tend to mind-wander more in others (Al-Balushi & Al-Harthy, 2015; McVay & Kane, 2012b). TUT rates thus seem to capture something meaningful about sustained attention during reading.

Most relevant to the present study, associations of comprehension with cognitive ability and motivation are partially mediated by mind-wandering propensity. McVay and Kane (2012b) had students read fiction and nonfiction texts, two of which included thought probes. Latent-variable analyses showed that students who mind-wandered more comprehended less. Moreover, students' WMC and attention abilities predicted comprehension scores, but partially through their associations with TUT rate. Unsworth and McMillan (2013) extended these findings to show that students' motivation to learn and their interest in the material predicted reading comprehension, and that they did so independently of WMC. Notably, these separate cognitive and noncognitive influences on comprehension were partially mediated by TUT rate during reading. Susceptibility to mind wandering thus appears to be one mechanism through which cognitive abilities and noncognitive inclinations influence reading comprehension.

Mind Wandering During Lectures

Laboratory studies. A closer analogue to the classroom comes from laboratory studies presenting thought probes within video-recorded lectures (e.g., Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2011; Szpunar, Khan, & Schacter, 2013). Despite variation in lecture lengths and topics, students consistently report TUTs at 30–40% of probes. Moreover, those who mind-wander more tend to recall less and, in studies examining temporal dynamics, TUTs increase across the lecture (Farley, Risko, & Kingstone, 2013; Risko et al., 2011; Risko, Buchanan, Medimorec, & Kingstone, 2013). Most relevant here, individual differences in mind wandering help explain normal variation in learning. Hollis and Was (2016) recorded students' TUTs during two video lectures from an online course and assessed their WMC and interest in the material. Consistent with the reading findings reviewed earlier, WMC and topic interest separately predicted TUT rate during online lectures, which, in turn, predicted lecture quiz scores and total course points. Moreover, TUT rate partially mediated the association between WMC and quiz score and fully mediated the association between topic interest and quiz score.

Classroom studies. Classroom investigations that probed students' thoughts during lectures, discussions, or other activities, have focused on contrasting the rates of different thought types across different contexts. Off-task thoughts unrelated to course content occur 10–35% of the time, with lower TUT rates during movies, demonstrations, and problem-solving activities (Locke & Jensen, 1974; Schoen, 1970) and higher TUT rates during lectures and student presentations (Bunce, Flens, & Neiles, 2010; Cameron & Giuntuli, 1972; Locke & Jensen, 1974). TUT rates are not trivial (15–25%), however, even during "active learning" contexts designed to maximally engage students (Geerlings, 1995; Shukor, 2005).

Only two classroom studies have examined the association between mind wandering and learning. Lindquist and McLean (2011) found that TUT rate during a single lecture correlated significantly but weakly with exam scores and final grades in the course (Kendall's $\tau \approx -.13$). In addition, students who confessed less interest in the lecture topic reported more TUTs ($\tau = -.11$). Wammes and colleagues (Wammes, Seli, Cheyne, Boucher, & Smilek, 2016b; see also Wammes, Boucher, Seli, Cheyne, & Smilek, 2016a) probed mind-wandering within most meetings of one course, asking students to distinguish intentional from unintentional TUTs at each probe, and quizzing them at the end of each meeting. Intentional TUT rates correlated with daily quiz scores ($r = -.21$), whereas unintentional TUT rates correlated with exam scores ($r = -.20$), but neither correlated with a preterm assessment of motivation to learn (rs = −.14 and .02, respectively). Both sets of classroom findings suggest that students who mind-wander more during lectures learn less, and both also show modest (sometimes nonsignificant) negative associations between TUTs and topic interest and motivation.

Limitations of Prior Educationally Relevant Mind-Wandering Research

Although these studies provided initial evidence concerning the relationship between mind wandering and educational outcomes, they are limited in two important ways: Most of the individual-differences studies reviewed above have (a) examined few cognitive factors and (b) inadequately assessed prior domain knowledge and topic interest.

Regarding (a), the previous literature as a whole has examined WMC, topic interest, prior knowledge, and motivation, but each study focused on only a smaller subset of these constructs. For example, in McVay and Kane's (2012b) reading study, the key individual-differences variable was WMC; similarly, in the Hollis and Was (2016) video-lecture study, only WMC and topic interest were assessed. In contrast, in the current study, we examined a broad range of individual-differences variables and systematically tested their respective contributions in regression and mediation analyses (**H1**, **H3**, & **H4**).

As for (b), we note two limitations in prior research. First, although domain knowledge is an important variable in any learning context, no studies of learning and mind wandering have pretested students' knowledge of the subject matter. Only two studies (Unsworth & McMillan, 2013; Wammes et al., 2016b) had subjects self-report their knowledge, but self-ratings are less ideal than objective assessment. Second, prior research assessed subjects' topic interest *after* (rather than *before*) the learning task was completed (Hollis & Was, 2016; Lindquist & McLean, 2011; Unsworth & McMillan, 2013; but see Wammes et al., 2016b). This practice confounds a true predictor variable (i.e., *prior* topic interest) with an outcome measure (i.e., topic interest *triggered* by the act of learning). In fact, the latter type of interest, called *situational interest*, has been studied in educational research as an important outcome variable (e.g., Hidi, 1990). Moreover, in one reading study that assessed interest *before* the texts were read, interest ratings did *not* predict TUT rate during reading (Fulmer, D'Mello, Strain, & Graesser, 2015). We addressed these limitations and contradictions by independently assessing subjects' knowledge and topic interest before the video lecture. In addition, we assessed, after the video, subjects' situational interest—their interest in the lecture and perceived utility of statistics triggered by the lecture—as one of the two key outcome variables of the study.

Is the Influence of Mind Wandering on Learning Always Negative?

Prior research, reviewed above, has established the negative effects of mind wandering on comprehension and learning. However, given that adults spend 25–50% of their waking time mind wandering (Kane et al., 2007; Killingsworth & Gilbert, 2010; Song & Wang, 2012), it might sometimes be adaptive. Indeed, Singer's pioneering work claimed that "positive-constructive daydreaming" contributed to everyday problem solving, creativity, and imagination (Singer, 1966; McMillan, Kaufman, & Singer, 2013). Similarly, Klinger's investigations of "fantasy" emphasized the role of mind-wandering in goal striving (Klinger, 1971, 2013). Furthermore, recent reviews and commentaries argue that mind wandering frequently occurs during routine activities, not only with little performance cost but also with considerable benefit of reflection, which may contribute to planning, delaying gratification, creativity, a sense of self, and deeper understanding of emotions (e.g., Immordino-Yang et al., 2012; Smallwood & Andrews-Hanna, 2013). In short, some forms of mind wandering—in the right contexts—may be beneficial.

The evidence for such beneficial effects of mind wandering, however, is severely limited in educationally relevant contexts. Several classroom studies (Locke & Jensen, 1974; Schoen, 1970; Shukor, 2005) collected open-ended thought reports from students and found nontrivial rates of thoughts that, while not focused on the here-and-now of the lecture, were also not entirely off-task (e.g., thoughts related to either course themes or reflected metacognitive evaluations of comprehension). Despite the presence of such off-task but topic-related thoughts, little is known about their cognitive consequences. For example, one reading study with older and younger adults (Frank, Nara, Zavagnin, Touron, & Kane, 2015) investigated whether theme-related or comprehension-related mind wandering would be associated with comprehension, but no clear relations were observed (in young adults, only comprehension-related thoughts were *negatively* correlated with comprehension scores). Jing, Szpunar, and Schacter (2016) investigated lecture-related mind wandering and learning from a video and reported *positive* associations (*r*s ≈ .45), but their sample was extremely underpowered for correlational analyses (n = 36, divided into 2 experimental groups). Such limited evidence suggests a need for further testing whether certain types of mind wandering (especially off-task but lecture-related thoughts) are indeed positively associated with posttest scores and situational interest (**H5**).

Note-Taking as an Intervention to Promote Focused Attention (for Some Students)

Another limitation of prior educationally relevant research on mind wandering is that few studies have considered how particular educational practices might influence mind wandering (for notable exceptions, however, see Jing et al., 2016; Szpunar et al., 2013). The current study examined note-taking as a potentially effective way to promote focused attention and thereby reduce mind wandering, at least among some students (**H2**).

Of course, students take notes primarily to create a written record for subsequent study. But, by actively taking notes, they may also pay more attention to, and thus better encode, the material (DiVesta & Gray, 1972). Indeed, a meta-analysis of 57 studies (Kobayashi, 2015) comparing students who did versus did not take notes on a reading or a lecture (without subsequent review) indicated a small encoding benefit (Cohen's *d* = .22, 95% confidence interval or CI [.17, .27]): Students assigned to take notes learn more than do those who do not take notes, perhaps because note-taking helps scaffold sustained attention. The meta-analysis also identified moderators of the encoding benefit (e.g., audiovisual versus text-only or audio-only materials), but did not consider individual differences.

A modest—and mixed—literature also suggests that cognitive and motivational factors may affect note-taking quality and benefits (for a review, see Bui & Myerson, 2014). WMC, for example, sometimes predicts students' note quality (e.g., Bui, Myerson, & Hale, 2013; Hadwin, Kirby, & Woodhouse, 1999; Kiewra & Benton, 1988; McIntyre, 1992), but sometimes does not (e.g., Bui et al., 2013; Peverly et al., 2007, 2013), perhaps because WMC effects depend on note-taking strategies (Bui & Myerson, 2014).

Individual-differences factors may also influence any benefits of note-taking, indicating individual-by-treatment interactions. Peper and Mayer (1978), for example, found that students with lower Scholastic Aptitude Test (SAT) scores scored higher on a near-transfer test when they *did not* take lecture notes, but higher on a far-transfer test when they *did*. Note-taking benefits may also interact with noncognitive factors. In a study where students learned from interesting or uninteresting passages, note-taking benefitted encoding of only low interest passages (Faber, Morris, & Liebermann, 2000). These findings suggest that note-taking might help students maintain attentional focus and limit their mind wandering, but that such note-taking benefits might be consequential for only some students (**H2**).

The Current Study

This study explored, in a diverse student sample (N = 182) drawn from two sites, individual-by-treatment interactions in the effectiveness of note-taking for reducing mind wandering during a video lecture on introductory statistics. We also tested whether any reduction in mind wandering would be associated with greater learning and situational interest in the lecture.

Treatment and Outcome Variables

Experimental manipulation. Half the students took notes during the lecture and half did not; all were subsequently tested on lecture content without reviewing notes. We focused our experimental manipulation on longhand note-taking because taking notes by hand has been shown to be more beneficial for conceptual learning than by computer, presumably because handwriting slows students down so they must put material into their own words rather than transcribing it verbatim (Mueller & Oppenheimer, 2014). Being slowed down and forced to paraphrase incoming information might also help students stay mentally on-task.

Outcome variables. One key outcome variable of the study was learning from a video lecture, indicated by a postvideo test (*posttest*) that assessed both factual and conceptual information and required multistep calculations. Note that, because we similarly pretested subjects' statistics knowledge before the video and included this pretest score in our regression models, our analyses of posttest performance reflected knowledge gained from the video.

Although learning is the main goal of most educational systems, teachers also hope to inspire students' interest in a way that motivates subsequent learning. Thus, the other outcome variable we examined was *situational interest* (e.g., Harackiewicz et al., 2008; Hidi, 1990; Kintsch, 1980), defined as momentary interest triggered by situational or environmental factors, here by viewing the video lecture. Immediately following the video, we assessed students' interests in, and perceived utility of, the lecture content. Of importance, and unlike several prior studies examining mind wandering and topic interest (e.g., Hollis & Was, 2016; Unsworth & McMillan, 2013; but see Fulmer et al., 2015), we assessed students' interest in math before the lecture and thus distinguished prelecture interest from the situational interest stimulated by lecture itself.

Five Primary Hypotheses

We tested a set of five hypotheses derived from the existing literature. Because no prior studies of mind wandering simultaneously included as wide a range of individual-differences measures as in the current study, we were not able to propose *a priori* the specific combinations of variables that would uniquely predict our dependent measures or interact with the note-taking manipulation. Thus, the current study was not entirely confirmatory in its approach and necessarily involved explorations of unique individual differences predictors and moderators. We instead tested more general hypotheses (**H1**–**H3**) that could still be theoretically informative and provide a basis for future confirmatory testing because we included only variables with theoretical or empirical support for their contribution to mind wandering and learning.

Our first hypothesis (**H1**), consisting of two subhypotheses, concerned the individual-differences variables that would uniquely predict mind-wandering rates during a video lecture:

H1a. *WMC and prior topic interest should independently predict TUT rates.*
H1b. *Additional measures, such as prior knowledge, incremental beliefs in math ability, learning confidence, and in-class media multitasking tendencies, should also predict TUT rates, perhaps with some doing so independently of the others.*

As reviewed earlier, most of the educationally relevant mind-wandering studies that examined individual differences assessed only WMC, topic interest, or both, as predictor variables (Hollis & Was, 2016; Unsworth & McMillan, 2013; Wammes et al., 2016b). These studies found that WMC significantly predicted mind-wandering rate, but the WMC–TUT association was modest at best, even at the level of latent variables. Also, it has been shown that topic interest uniquely predicts TUT rate during learning, but, in some cases (e.g., Hollis & Was, 2016; Lindquist & McLean, 2011; Unsworth & McMillan, 2013), it was assessed only after (or combined both before and after) the reading or lecture was completed, making the measure an ambiguous mix of prior and situational interest. We therefore tested **H1a** in the current study.

H1b concerned the unique predictive power of various variables that prior research has not examined optimally or even not at all. For example, prior knowledge is well known to facilitate new learning (e.g., Ambrose, Bridges, DiPietro, Lovett, & Norman, 2010), but, because few mind-wandering studies have objectively assessed knowledge, it is not yet clear how prior knowledge affects attention toward (or away from) task-relevant information. We thus pretested subjects for statistical knowledge and asked about their prior math coursework. We also assessed, prior to the video lecture, two belief/motivation-related variables that we believed might predict during-lecture mind wandering: (a) students' beliefs that math ability is modifiable (growth vs. fixed mindset) and (b) confidence in learning from the lecture.

Finally, we assessed the academic habit of media multitasking. No prior studies have tested its association with TUTs, but students who engage in more multitasking tend to be worse in their attention-test performance (e.g., Cain, Leonard, Gabrieli, & Finn, 2016; Ophir, Nass, & Wagner, 2009). Classroom multitasking propensity, then, may predict TUTs during, and learning from, on-line lectures. This measure was particularly of interest because, in the current study, none of the subjects were allowed to do any media multitasking during the experiment (e.g., students had no access to their smartphones or tablets). In such a situation, which mimicked a typical learning context, we hypothesized that propensity toward in-class media multitasking might uniquely (and positively) predict mind-wandering rate during the lecture.

The second hypothesis (**H2**) concerned the effects of note-taking on mind-wandering:

H2a. *Active note-taking during the lecture should reduce TUT rates and thus increase posttest scores and situational interest.*

H2b. *If note-taking moderates the effects of individual-differences variables on TUT rates, it should do so by reducing the magnitudes of these individual-differences effects.*

H2c. *The effectiveness of note-taking in reducing mind wandering likely depends on levels of compliance and engagement and, hence, note-taking quality/quantity should predict*

TUT rates.

If the encoding benefit of note-taking (Kobayashi, 2005) reflects an improvement in learning via the scaffolding of attention, then note-taking should reduce mind wandering overall (**H2a**). It should also particularly benefit students who otherwise struggle to maintain focus, thus minimizing the individual differences in mind wandering, learning, and situational interest that would normally be seen in the absence of note-taking (**H2b**). Of course, not every student assigned to take notes will likely do so effectively; that is, we cannot control the "dosage" of this intervention because we cannot force students to take notes of high quality and quantity. Thus, we hypothesized that the benefits of note-taking might be most clearly seen by examining the quality and quantity of notes taken by students in the note-taking condition (**H2c**).

The third hypothesis (**H3**) concerned whether mind wandering is a unique predictor of the two outcome variables above and beyond other individual differences variables:

H3. *TUT rates should uniquely predict both posttest scores and situational interest above and beyond other relevant individual differences*.

In laboratory and classroom studies, TUT rates tend to predict learning outcomes as strongly as, or more strongly than, cognitive ability, motivation, or interest measures do (Hollis & Was, 2016; Lindquist & McLean, 2011; McVay & Kane, 2012b; Unsworth & McMillan, 2013; Wammes et al., 2016b). Thus, we rigorously tested the hypothesis (**H3**) that TUT rate would be an independent predictor of posttest scores and situational interest, not only by including ability, habit, and belief measures in our regression models, but also by including pretest score as a predictor of posttest scores and prior math interest as a predictor of situational interest.

The fourth hypothesis (**H4**) concerned whether mind wandering rate mediates the associations between individual differences variables and the outcome measures:

H4. *The predictive power of individual-differences variables for posttest and situational* *interest outcomes should be significantly mediated by TUT rates during the lecture.*

Prior individual-differences studies of mind wandering in educational contexts indicate that TUT rate not only accounts for unique variance in learning and comprehension but also significantly mediates the associations between some individual-difference variables and learning outcomes. Thus, we hypothesized that mind wandering is one mechanism through which the individual-differences predictors affect learning and deriving interest from a lecture.

The last hypothesis (**H5**) concerned potential benefits of mind wandering:

H5. *Lecture-related TUTs, but not comprehension-related TUTs, will positively predict posttest scores and situational interest.*

As reviewed earlier, students sometimes report off-task thoughts that are not focused on the present moment but are actually related to task content or their comprehension of that content (Locke & Jensen, 1974; McVay & Kane, 2012b; Schoen, 1970). Metacomprehension thoughts do not consistently predict outcomes (Frank et al., 2015), perhaps because they reflect active engagement for some subjects but struggles to understand in others. In contrast, on-topic, lecture-related mind-wandering may be associated with better learning (Jing et al., 2016), perhaps because it helps students connect current material with prior knowledge, although this latter finding came from a small sample. We thus tested the hypothesis that lecture-related (topic-related), but not comprehension-related, mind wandering is systematically related to the two outcome measures of the study, posttest scores and situational interest (**H5**).

Methods

Below we report how we determined our sample size and all data exclusions, manipulations, and measures in the study (Simmons, Nelson, & Simonsohn, 2012). Moreover, the stimulus materials used in the current study (e.g., the video lecture, the pretest/posttest statistics questions) as well as the files containing anonymized data are available for downloading at the following URL: [URL here].

Subjects

We tested 200 undergraduates, 100 each from the University of Colorado Boulder (hereafter Site A), a flagship state university in the Western U.S., and the University of North Carolina at Greensboro (hereafter Site B), a comprehensive state university and minority-serving institution for African-American students in the Southeastern U.S. All subjects were 18–35 years old and participated for partial fulfillment of a research requirement for an introductory psychology course. Half the subjects at each site were pseudorandomly assigned to either a note-taking or no-notes condition, via a prerandomized sequence sheet stored in each testing room; all subjects within a session were assigned to the same condition. We tested, but did not retain or analyze the data from, additional subjects who indicated on the math/statistics background questionnaire (see below) that they had taken a formal course on statistics.

Power Considerations and the Rationale for Multisite Data Collection

Our study of mind wandering was unique in both assessing a large number of individual-differences variables and testing an experimental note-taking intervention (and, thus, assessing individual-by-treatment interactions in the prediction of mind wandering, learning, or situational interest). Thus, we had little information from prior studies that we could use to conduct a formal *a priori* power analysis. We knew, however, that many of the pairwise associations of interest to us (e.g., mind wandering and learning) were in the range of r = .20 (e.g., Al-Balushi & Al-Harthy, 2015; Hollis & Was, 2016; McVay & Kane, 2012; Risko et al., 2013; Schooler et al., 2004; Unsworth & McMillan, 2013; Wammes et al., 2016). We therefore sought a sample size of 200 to ensure 80% power to detect most anticipated correlations and to elicit reasonably stable effect-size estimates of anticipated correlations of .20 (i.e., close to 80% power to reach the critical point of stability within a corridor width of .10 [required *N* = 238] and well above 80% power for a corridor width of .15 [required *N* = 104]; Schönbrodt & Perugini, 2013).

Our primary motivation in collecting data from two research sites was to expand the diversity of our student sample and ensure substantial variability in the individual-differences predictor and outcome measures. In fact, as summarized in Appendix A, the mean scores differed significantly between the two sites in all outcome variables and two of the individual-differences predictor variables. Although we assumed that key results would be sufficiently robust across sites, such site differences require an explicit test of generalizability. We present supporting evidence at the end of the Results section, although cross-university comparisons were not our primary interest in conducting this multisite study.

Procedure, Materials, and Equipment

Materials and equipment. Each subject sat in front of a Mac Mini computer with an Acer 22-inch LED-LCD monitor and, during the video portion of the procedure, wore Koss UR-20 headphones. We also provided each subject with a Sharp EL243SB calculator. A Marpac DOHM-DS white noise machine ran at the low setting throughout each session. Subjects in the note-taking condition were provided with a three-ring notebook filled with lined loose-leaf paper.

Overall procedure. Testing rooms accommodated up to 3 subjects per session at Site A and up to 4 per session at Site B. The experimenter remained throughout the session and read aloud all onscreen instructions. After each questionnaire or task, subjects in group sessions waited until all others were finished before the experimenter read the next task's instructions and allowed subjects to move on. Most sessions lasted 90–120 min. Below, we describe the measures and tasks administered in 5 stages, in order of appearance.

Stage 1: Individual Differences Assessment

After informed consent, subjects completed a series of individual differences measures, all administered via computer.

Symmetry span. We measured WMC with an automated symmetry span test (adapted from Redick et al., 2012). Each trial presented 2–5 processing items in alternation with memory items. Each processing item presented a black-and-white pattern within an 8 x 8 matrix that subjects judged (via mouse-click) as either vertically symmetrical or asymmetrical; each memory item was a 4 x 4 matrix with one to-be-remembered square in red, presented for 650 ms. Following each trial sequence, a recall screen appeared that presented a blank 4 x 4 matrix. Subjects recalled the 2–5 red squares from that trial by clicking on their locations in the blank matrix in serial order (subjects could click on a "blank" button for any item in sequence that they could not recall). Subjects completed two trials at each set size, presented in randomized order, with the constraint that all set sizes were presented once before a set size repeated. Prior to the real trials of the task, subjects first practiced the memory portion alone (one trial each of set sizes 2 and 3), then practiced the processing portion alone (two trials each of set sizes 2–5), and then practiced two combined trials (one each of set size 2 and 3). Processing-only practice was additionally used to establish the RT cut-off for processing stimuli in the real task: If any symmetry-decision display exceeded the processing-practice mean RT plus 2.5 SDs, it disappeared and counted as a processing error. The dependent measure for the symmetry span task was the proportion of 28 memory items recalled in the correct serial position.

Questionnaires. Subjects then completed four questionnaires in the order presented below. Instructions indicated that subjects had the right to skip any item, and subjects also saw a confirmation screen for any skipped question without a selected answer, on which they could confirm the skip or go back to answer the question. For all the questionnaires administered before and after the video (except where specified below), subjects saw all items in the same randomized order, untimed, one at a time.[1] For most of the questionnaires, response options were presented next to an empty box, and subjects responded by mouse-clicking on the box next to the chosen option. For the questionnaires using a 1–5 Likert scale, we averaged the numbers corresponding to each response after reverse scoring appropriate items. For some questionnaires, however, some items were not included in the averaging, as noted below.

Math/statistics background questionnaire. This questionnaire asked open-ended numerical questions: (a) one about total college credits completed, (b) three about the number of high school and college courses completed in mathematics, and then (c) one about courses in statistics. For this questionnaire, we used the number of high-school math courses completed as a predictor variable, after setting all values <3 to 3 and values >6 to 6 (given Colorado and North Carolina state high school graduation requirements, values <3 and >6 were unrealistic).

Note-taking behavior questionnaire. Eleven items (adapted from online note-taking-habit surveys) asked about the student's typical note-taking practices and skills (e.g., *I have difficulty putting class notes in my own words; I know what is the 'important stuff' to write down and what are the cues that this is important stuff*). The 1–5 response scale was labeled, *Never, Rarely, Sometimes, Often, Always*. For this questionnaire, two items were not included in the score: *I take notes in class on paper or in a notebook* and *I take notes in class with a laptop or tablet*. This is because, unlike the other items, they did not assess self-reported note-taking ability or skill.[2]

---

[1] Three of the questionnaires also included a catch question (e.g., *I write my notes by alternating between Dutch and Portuguese*) to assess random or inattentive responding (Maniaci & Rogge, 2014). We replaced all data from the one subject who missed more than one catch question.

[2] Self-reported tendencies to take class notes *on paper* (M = 4.53, SD = 0.76) or *on a laptop* (M = 1.86, SD = 1.01), did not correlate significantly with TUT rate, $rs(180) = .05$, $p = .484$, and $-.07$, $p = .347$, respectively.

Multitasking questionnaire. Three items asked subjects to report how much they engaged in in-class: (a) "visual" media multitasking (e.g., texting, emailing, web surfing); (b) doodling on paper; and (c) daydreaming. The 1–5 response scale was labeled, *Never, Rarely, Sometimes, Often, Always*. For this questionnaire, we used only the first question, about "visual" media multitasking in class, for our primary analyses.[3]

Mathematics/statistics interest and beliefs questionnaire. This questionnaire combined two types of items: (a) 9 items (drawn or modified from Eccles & Wigfield, 1995; Linnenbrink-Garcia et al., 2010; Pintrich & DeGroot, 1990) about students' interest in, and perceived utility of, mathematics/statistics (e.g., *I like mathematics/statistics*; *It is important for me to be a person who reasons mathematically/statistically*; *Mathematics/statistics will be useful for me later in life*); and (b) 2 items about belief in the incremental versus fixed nature of mathematics/statistics ability (adapted from Dweck, 1999). For both, the 1–5 response scale was labeled, *Strongly Disagree, Somewhat Disagree, Neither Agree or Disagree, Somewhat Agree, Strongly Agree*. For these questionnaires, we calculated two scores—one for math interest (averaged across 9 items) and one for incremental beliefs about math ability (averaged across 2 items).

Stage 2: Prevideo Statistics Test (Pretest) and Learning Confidence Assessment

Subjects next completed the prevideo statistics test (hereafter, *pretest*).

Pretest. The pretest consisted of 10 multiple-choice questions, administered with no time limit, and with the provided calculator. We had piloted questions to yield low pretest scores, to show a mean increase in accuracy from pretest to posttest, and to elicit good item-total

[3] Reports of doodling in class (M = 2.63; SD = 1.06) correlated significantly with TUT rate during the video, $r(180) = .23$, $p = .002$, but did not correlate with posttest performance, $r(180) = -.01$, $p = .899$, or situational interest, $r(180) = -.05$, $p = .479$. Classroom daydreaming (M = 3.37; SD = 0.86) correlated positively with in-lecture mind wandering rate, $r(180) = .30$, $p < .001$. Although classroom daydreaming was not related to posttest scores, $r(180) = -.07$, $p = .340$, it was negatively correlated with situational interest, $r(180) = -.18$, $p = .016$.

correlations. Each question was followed by 6 or 7 answer choices labeled by the letters A–F or A–G, with a box to the left of each choice. Subjects responded to each question by mouse-clicking the appropriate box. They were told to guess if they did not know the answer because there was no penalty for guessing and, upon making their selection, to mouse-click on a "submit" button onscreen. After each response, a pop-up screen asked subjects about their confidence in their answer. Subjects indicated via mouse-click whether they: (a) had to guess with little confidence; (b) had to guess but with some confidence; or (c) knew the answer with high confidence.[4] The dependent measure was the proportion correct out of 10 questions.

Learning confidence assessment. Subjects were then asked two questions about their pretest performance, first to estimate how many of the 10 questions they got correct, and second to predict how many of those same questions they would be able to answer correctly after watching the video lecture covering those topics. Because the 10-item questions were intentionally made challenging and because we excluded any participants who indicated some systematic prior exposure to statistics, subjects' responses to the first question were mostly low. We thus used the latter measure—how many questions they would answer correctly after the lecture—as the measure of learning confidence.

Stage 3: Video Lecture

Subjects then watched a realistic video lecture (while listening through headphones), during which we assessed their mind wandering frequency (TUT rate) using thought probes.

[4] Subjects reported knowing the answer with high confidence on 31% of their correct answers and 10% of their wrong answers, guessing the answer with some confidence on 26% of their correct answers and 30% of their wrong answers, and guessing the answer with little confidence on 42% of their correct answers and 59% of their wrong answers. Only 3 subjects indicated knowing the answer with high confidence to more than 3 questions while getting those 3 questions correct. These results from the confidence ratings suggest that these pretest items were difficult enough for our subjects.

Lecture content. The video lecture (recorded via Camtasia Studio 8 software; presented via an E-Prime 2.0 program) consisted of a PowerPoint presentation showing images and text accompanied by audio narration. The lecture was divided into 31 video segments presented seamlessly. The first segment was 5 min long and each remaining segment was 1:08–1:51 min in length. The entire video, not including thought probes, lasted 52 min.

The video lecture taught students to understand and calculate the SD for a set of scores. It began with a brief introduction to the everyday utility of statistics and some definitions (e.g., populations, parameters, samples, and statistics). Then, using the example of pretesting a group of five high school students on a verbal SAT subtest, the video defined and explained descriptive statistics, focusing on creating and interpreting frequency distributions. The lecture then explained central tendency and different definitions of "average," including mode, median, and mean, using the SAT subtest example to calculate each. The mean was discussed in terms of its calculation, its mathematical formula (e.g., introducing summation and its symbol, Σ), and its conceptual role as the balance point of a distribution.

Continuing with the verbal SAT example, the lecture then explained variability and how it might differ across distributions despite identical means; it then defined and explained range and SD. Following this conceptual introduction to the SD, the video used the SAT example to explain its calculation incrementally, from calculating the mean, to the deviation scores, to the squared deviation scores (and the reason for these), to the sum of squares, to the variance, to the SD. To reinforce the SD calculation, the video then worked through a second example with a sample of 10 SAT subtest scores. The video ended by considering how to interpret a given SAT score when one knows the mean and SD of the typical SAT distribution.

Video instructions. Before the video, the experimenter read onscreen instructions encouraging subjects to try to learn as much as they could; these instructions told subjects that the video was a narrated PowerPoint presentation, asked them to learn as much as they could, and noted that they could not pause or rewind the lecture. The experimenter and onscreen instructions then introduced the note-taking manipulation, asking subjects in the note-taking condition to use the provided notebook to take notes on the lecture, just as they would in a class in which they expected to be tested on the material. The experimenter asked subjects in the no-notes condition to listen to the lecture but not write anything down, just as they would in a real class in which they did not take notes but were expecting to be tested later.

Mind-wandering probes and instructions. All subjects then learned about the thought probes that would appear throughout the video lecture, each as a green screen, to assess their thought content. The experimenter explained that, while watching the video, they might occasionally find themselves thinking of something else, and that the study was interested in the types of things that people think about during lectures and other learning contexts:

> *In order to examine this, the computer will periodically interrupt the video to ask you what you were *just* thinking about, at the very instant before the computer asked you. It is perfectly normal to occasionally find yourself thinking about things that are not related to the ongoing task you're doing. Every now and then, the computer will present you with a green screen that has several categories of things that people might be momentarily thinking about during a video lecture like this one. Whenever you're asked by the computer at this green screen, please try your best to honestly take stock of what your thoughts had *just* been about that instant. Then please choose a category that best describes what your thoughts were about.*

The experimenter then explained the seven response options:

1. ***On-task on the lecture***, for thoughts about what was being discussed in the video at that time;
2. ***Lecture-related ideas***, for thoughts about some aspect of the lecture topic, but not what was being presented in the video at that moment;
3. ***How well I'm understanding the lecture***, for evaluative thoughts about comprehending (or not) what was being presented on-screen;
4. ***Everyday personal concerns***, for thoughts about normal everyday things, life concerns, or personal worries;

5. ***Daydreams***, for fantasies or unrealistic thoughts;
6. ***Current state of being***, for thoughts about one's current physical, psychological, or emotional state (e.g., thinking about being sleepy, hungry, or fascinated); and
7. ***Other***, for any thoughts not fitting into the other categories.

The probes appearing during the video displayed each number and italicized label above, all appearing below the question, *What were you just thinking about?* Subjects responded by pressing the numerical key corresponding to their choice. Further instructions emphasized responding to probes based on immediately preceding thoughts:

> *Please remember to always respond to what you had *just* been thinking about, as if the green screen is like a flashbulb or photograph that captures and preserves a moment in time. Please do not try to reconstruct what you'd been thinking about over the last few seconds or minutes.*

During the video, each subject saw 20 probes. Probes were randomized for each subject to appear between the video segments, with the constraint that no more than three segments in a row were probed. Subjects could take as much time as needed to respond. We presented probes randomly with constraints to make contact with prior studies assessing mind wandering during laboratory tasks (including reading).[5]

Scoring of TUT rate. Mind wandering scores represent the proportion of the 20 probes on which subjects indicated a TUT by selecting response choices 4–7 above. In targeted analyses for **H5**, we also examined proportions of off-task but lecture-related thoughts (choice 2) and comprehension-evaluative thoughts (choice 3).

Postvideo questions. After the video, the experimenter collected notebooks from the note-taking subjects, so that all posttests were completed "closed-book." The computer then presented subjects with two metacognitive questions for a numerical response: (a) *What*

[5] This random placement of probes, however, likely introduced noise in the individual-differences measurement if different sections of the video elicited more or less mind wandering. Thus, we may have underestimated the effect sizes for TUTs' association with other predictor variables or outcome measures.

percentage of the material and information in the lecture do you think you were able to understand and remember?; and (b) *If you are now given the same 10 questions that you answered before the lecture, how many of them do you think you can answer correctly?* We did not include these retrospective metacognitive items in the analyses reported below.

Stage 4: Assessment of Outcome Measures (Situational Interest and Posttest)

After the video lecture, subjects completed two outcome measures: the situational interest questionnaire and the postvideo statistics test (hereafter, *posttest*).

Situational interest. This questionnaire consisted of 10 items (modified from Linnenbrink-Garcia et al., 2010), assessing interest in the lecture and in statistics (e.g., *I found the content of this video lecture personally meaningful*; *To be honest, I just don't find statistics interesting*; *I think what I learned from this video lecture is useful for me to know*"). The 1–5 response scale was labeled, *Strongly Disagree, Somewhat Disagree, Neither Agree or Disagree, Somewhat Agree, Strongly Agree*. The dependent measure was the average of the two subscales about the lecture itself ("interest in the lecture" and "utility of the lecture," of 3 and 4 items, respectively).[6]

We presented this questionnaire differently from the others: To equate the retention interval between the video material and the upcoming posttest for all subjects, we fixed the presentation time for each questionnaire item. Each item appeared for 4.5 s and then the screen background changed from white to yellow, indicating that subjects now had 5 s to type their numerical response (the program did not accept responses during the initial 4-s window). Thus, regardless of when subjects typed their response, each item remained onscreen for 9.5 s.

[6] We did not analyze the 3-item subscale for subjects' interest in statistics as a field, primarily because we did not expect viewing a single lecture video on statistics to change a student's inherent interest in statistics. Nevertheless, this subscale (M = 2.58, SD = 0.86) significantly correlated with TUT rate, $r(180) = -.44$, $p < .001$ and with posttest scores, $r(180) = .23$, $p = .001$.

Posttest. Instructions for the posttest followed immediately after the situational interest questionnaire. Subjects completed Parts 1–3 at their own pace (with the calculator, as needed), but all subjects in a session had to complete the current part before anyone could move on to the next. Part 1 was exactly the same as the 10-item pretest questions but without the confidence/guessing probes, whereas Parts 2 and 3 required subjects to calculate a SD, the focus of the lecture. Part 2 asked subjects to calculate the SD of four scores (*1, 3, 5, 7*). In addition to entering their final answer, subjects showed all their work for all calculation steps on a provided piece of scratch paper. Part 3 required subjects to calculate the SD of five scores (*2, 3, 4, 7, 9*), but here the process was divided into the five instructed steps: (a) computing the mean of the scores; (b) computing the deviation scores; (c) computing the sum of squares; (d) computing the variance; and, finally, (e) computing the SD. As in Part 2, subjects showed their work and provided their answer, but here on a separate piece of scratch paper per step.

The dependent measure was the mean of the standardized scores (*z* scores) for Part 1, Part 2, and Part 3, each weighted equally. Like the pretest score, the posttest Part 1 score reflected proportion correct out of 10 questions. For the scoring of Parts 2 and 3, two independent coders evaluated the calculations provided on scratch paper to award partial credit (interrater reliability, as indicated by intraclass correlation, was .91 for Part 2 and .93 for Part 3). One point was given for each step correctly performed with a total possible score of five points. If subjects indicated the correct procedure for a step but miscalculated the final outcome for that step, 0.5 points were awarded. Also, once a step was miscalculated, any subsequent step that was correctly performed also only received 0.5 points. For example, even if, during the first step of Part 3, a subject incorrectly calculated the mean to be 4.00, they still received 0.5 points if they showed the correct procedure (i.e., "(2 + 3 + 4 + 7 + 9) / 5"). Even though Part 2 did not provide separate sheets of scratch paper for each step, this same scoring rubric was used.

Stage 5: Demographic Questionnaire

Subjects completed a demographic questionnaire about gender, age, ethnicity, race, and college major. We administered this questionnaire at the end to minimize any potential negative effects of stereotype threat (e.g., Spencer, Steele, & Quinn, 1999; Steele, 1997) that might arise from indicating gender (and possibly race/ethnicity), given the mathematical focus of the study.

Coding of Note-Taking Quality/Quantity

For subjects in the note-taking condition, we created a combined quality/quantity score for notes taken about the lecture, following Bui et al. (2013). We identified 31 key elements of the lecture and weighted them equally for a possible maximum score of 31 points. These elements were broken up into three sections of the lecture: (a) terms defined in the lecture (e.g., mean, variance; 20 points); (b) description of the five steps in calculating SD (5 points); and (c) the second five-step SD-calculation example presented at the end of the lecture (5 points), plus an additional point (1 point) if the subject noted the original problem (some subjects wrote the initial raw scores only and did not continue following through the example). For the first two sections of the lecture (a & b), we awarded subjects 1 point if they mentioned each of the key elements. For the final section (c), where the lecture worked through a concrete example of calculating the SD, the subject had to show their work for a step to receive the point.

Two independent researchers coded lecture notes. Reliability was high with a between-coder correlation of $r = .97$. The majority of discrepancies arose by one coder failing to recognize the presence of a key element in the notes. All discrepancies were resolved by having both coders read through notes again and agree on the source error and the appropriate solution.

Data Analyses

Data loss. Due to a technical issue, we lost the symmetry span data from one subject in the no-notes condition. In addition, we excluded data from 17 subjects from analyses because: (a) their accuracy in the processing part of the symmetry span test was below our 75% criterion (3 and 6 subjects, respectively, in the note-taking and no-notes conditions); (b) their accuracy in the pretest was higher than .60 (2 and 5 subjects); or (c) their performance estimates used to compute scores for learning confidence reflected values larger than 10 (1 subject in the note-taking condition). The final sample thus consisted of 182 subjects (*n*s = 94 and 88, respectively).

Demographic information. Table 1 summarizes the demographic characteristics of the final 182 subjects (collapsed across Sites A and B). As indicated in the table, the two groups did not differ on any of the key demographic variables.

Table 1. *Summary of demographic variables (collapsed across the two research sites)*

	Overall	Notes Condition		Condition Differences	
		No-Notes	Note-Taking		
	(N = 182)	(n = 88)	(n = 94)	χ^2	***p***
Sex				0.59	0.44
Male	47 (25.8%)	25 (28.4%)	22 (23.4%)		
Female	135 (74.2%)	63 (71.6%)	72 (76.6%)		
Race				0.54	0.91
White	121 (66.5%)	58 (65.9%)	63 (67.0%)		
Black	29 (15.9%)	14 (15.9%)	15 (16.0%)		
Multiracial	16 (8.8%)	9 (10.2%)	7 (7.5%)		
Other [a]	16 (8.8%)	7 (8.0%)	9 (9.6%)		
Ethnicity				0.00	0.97
Latino/Hispanic	25 (13.7%)	12 (13.6%)	13 (13.8%)		

Note. [a] This category ("Other") included: Asian, Native American/Alaskan Native, and Native Hawaiian/Pacific Islander.

General data-analysis plan. Our sample was drawn from two universities. Given that the availability of only two level-2 clusters (2 research sites) precludes multilevel modeling analyses (Hox, 1998; Hox, Moerbeek, Kluytmans, & van den Schoot, 2014), and given that we entertained no hypotheses regarding research site (which is likely confounded with dozens of unmeasured variables), we did not include research site (university) as a variable in the primary analyses reported below. As noted earlier, however, we conducted and will briefly discuss supplementary analyses that included research sites as a fixed-effect variable and examined the extent to which the key results generalized across sites (see Appendixes A, B, and C).

For all multiple regression analyses reported in this article, we standardized all of the continuous variables (both outcome and predictor variables) across the entire sample. The only exception concerns the site-by-site regression results reported in the online supplementary materials (Tables S1, S2, & S3), for which standardization was done for each site. We conducted meditational analyses and tested the indirect effects, using the *MEDIATE* macro for the SPSS statistical package (Hayes & Preacher, 2014). We adopted an alpha level of .05 throughout.

Results

Preliminary Analyses

TUT rates. Figure 1 presents the mean rates (proportions) of each broad category of thought report. Of most importance, TUT rates (i.e., proportions of probes on which subjects indicated off-task mind wandering) represented nearly half of subjects' thought reports, comprised of everyday personal concerns (M = .09 of all responses), daydreams (M = .10), thoughts about current state (M = .21), and other thoughts (M = .05). As indicated in Table 2, despite mind wandering being frequent overall, we found considerable interindividual variation, with TUT rates of approximately .20–.70 within one SD of the mean.

We also examined the time course of mind wandering by dividing the 20 thought probes into first, second, third, and fourth quarters for each subject. TUT rates increased across quarters (Ms = .39, .47, .49, and .46, respectively), yielding a significant main effect of quarter,

$F(3, 543) = 8.72$, $MSE = 0.040$, $p < .001$, $\eta_p^2 = .046$, a significant linear trend, $F(1, 181) = 14.01$, $MSE = .041$, $p < .001$, $\eta_p^2 = .072$, and a significant quadratic trend, $F(1, 181) = 10.45$, $MSE = .045$, $p = .001$, $\eta_p^2 = .055$. The latter indicates that mind wandering increased primarily between the first and second quarters, and then changed little.

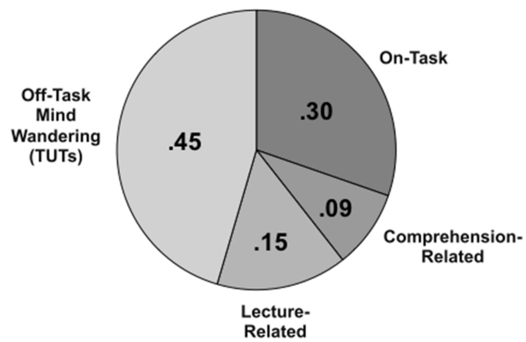

Figure 1. Mean proportion of each thought type reported across 20 probes per subject (these are across-subject *M*s, rather than response totals, so they do not necessarily add up to 1.00).

Effects of note-taking. Table 2 presents the descriptive statistics for all measures. Subjects reported off-task thinking to about half of the probes in both the note-taking and no-notes conditions. Although the effects of note-taking will be discussed more formally in subsequent regression models (**H2**), this observation indicates that simply providing students with the opportunity to take notes did not reduce off-task thinking at the overall group level (i.e., no main effect of note-taking). Indeed, note-taking did not appear to have much direct influence on the learning (posttest) or situational interest outcomes, either. As shown in Table 2, however, subjects did learn from the video, with mean pretest accuracy rates of ~.25 and mean posttest (Part 1) accuracy rates of ~.47 across experimental conditions.

Zero-order correlations. Correlations among the individual differences measures and the outcome measures are presented in Table 3. These correlations indicate that some of the measures were significantly associated with TUT rate. Specifically, students who reported more media multitasking in class had a higher in-lecture TUT rate, $r(180) = .23$, whereas students with

Table 2. *Descriptive statistics for outcome and predictor measures by note-taking condition*

	Notes Condition								Condition Differences	
	No-Notes (n = 88)				Note-Taking (n = 94)					
Measure	***M***	***SD***	**Min**	**Max**	***M***	***SD***	**Min**	**Max**	***t***	***p***
Outcomes										
Mind Wandering Rate	0.46	0.26	0.00	1.00	0.44	0.25	0.00	1.00	-0.54	.591
Posttest (Z-score aggregate)	-0.11	0.94	-1.62	1.55	0.10	1.05	-1.80	1.58	1.47	.140
Posttest (Part 1)	0.45	0.19	0.10	0.90	0.49	0.23	0.00	1.00	1.31	.193
Posttest (Part 2)	2.22	1.67	0.00	5.00	2.66	1.83	0.00	5.00	1.71	.089
Posttest (Part 3)	2.83	1.49	0.00	5.00	3.04	1.61	0.00	5.00	0.90	.368
Situational Interest	2.71	0.67	1.00	4.33	2.79	0.76	1.00	4.50	0.76	.447
Predictors										
Pretest	0.24	0.13	0.00	0.60	0.27	0.14	0.00	0.50	1.36	.175
Number of Math Courses	4.32	0.77	3.00	6.00	4.48	0.85	3.00	6.00	1.34	.183
Math Interest	2.97	0.97	1.00	4.78	2.86	0.97	1.11	4.56	-0.73	.469
Learning Confidence	6.35	1.98	1.00	10.00	5.91	2.12	0.00	10.00	-1.44	.152
Incremental Beliefs	3.49	0.88	1.50	5.00	3.64	0.97	1.00	5.00	1.05	.295
Media Multitasking Habits	2.60	0.97	1.00	5.00	2.71	1.08	1.00	5.00	0.73	.468
Note-Taking Habits	3.61	0.42	2.44	4.44	3.66	0.44	2.33	4.56	0.72	.473
WMC	0.67	0.18	0.11	1.00	0.65	0.20	0.09	1.00	-0.90	.371

Note. WMC = working memory capacity. Degrees of freedom (*df*s) for the *t* tests = 180.

Table 3. *Bivariate correlations among all predictor and outcome variables (with internal reliability estimates indicated on the diagonal)*

Measure	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Outcomes											
1. Mind Wandering Rate	*(.85)*										
2. Posttest	**-.48**	*(.68)*									
3. Situational Interest	**-.56**	**.32**	*(.81)*								
Predictors											
4. Pretest	**-.23**	**.39**	.11	*(.09)*							
5. Number of Math Courses	-.12	.04	.03	.07	–						
6. Math Interest	**-.27**	**.28**	**.39**	**.17**	**.20**	*(.93)*					
7. Learning Confidence	-.09	**.15**	**.15**	**.17**	.03	**.34**	–				
8. Incremental Beliefs	**-.24**	.12	**.21**	.14	.08	**.28**	.07	*(.64)*			
9. Media Multitasking Habits	**.23**	-.09	**-.19**	-.01	.05	-.11	-.03	**-.23**	–		
10. Note-Taking Habits	-.13	.06	.11	.04	**.16**	.09	.09	.03	-.10	*(.65)*	
11. WMC	-.12	**.19**	-.04	**.18**	.10	.06	.06	.13	.02	**-.14**	*(.60)*

Note: Degress of freedom *(dfs)* for the correlations = 180. Significant correlations are printed in bold. WMC = working memory capacity. Reliabilities, where applicable, appear italicized and within parentheses on the diagonal (Posttest = coefficient α across 3 subparts; Situational Interest = α across all items; Mind Wandering = α across 4 quartiles of the video lecture; Pre-Test = α across all items; Math Interest = α across all items; Incremental Beliefs = α across both items; Note-Taking Habits = α measured across all items excluding the "on paper" and "on laptop" items; WMC = α across all set) lower pretest scores, less prior interest in math/statistics, and less incremental beliefs about math intelligence reported more off-task thoughts, *rs*(180) = −.23, −.27, and −.24, respectively. WMC did not significantly predict TUT rate in this study, *r*(180) = −.12, *p* = .107, but it did predict both pretest and posttest scores, *rs*(180) = .18 and .19, respectively. Of primary importance, subjects who mind-wandered more during the lecture scored worse on the posttest, *r*(180) = −.48, *p* < .001, and found it less interesting, *r*(180) = −.56, *p* < .001. These correlations provide some preliminary support for the hypothesis (**H4**) that TUT rate may serve as a mediator of the relationship between some individual differences measures and the outcome variables.

Note that some measures had modest reliabilities (indicated in the diagonal in Table 3), which likely limited somewhat their correlations with other measures (e.g., incremental beliefs, note-taking habits, and WMC; α = .64, .65, and .60, respectively). Pretest score had particularly low reliability (α = .06) in part due to our exclusion of subjects who scored high (>60%) on this measure, and also because subjects guessed on most questions, as explained in Footnote 4. That said, the alpha seemed to have underestimated the true reliability here, as pretest scores correlated with (and predicted unique variance in) TUT rate and posttest scores.

H1 and H2: Individual-Differences Predictors of, and Note-Taking Effects on, Mind Wandering

Data analysis plan. We first assessed the individual-differences predictors of TUT rate (**H1**) and the main and moderating effects of note-taking (**H2**) by running a single regression model (**H2c** was tested in a separate model, as noted later). TUT rate served as the dependent variable, and the model included three types of predictor variables: (a) individual differences measures, (b) the note-taking variable (contrast-coded: −0.5 for no-notes; +0.5 for note-taking), and (c) the first-order interaction terms involving (a) and (b). We had no *a priori* predictions for higher-order interactions (e.g., interactions between multiple individual-differences variables), so we did not include them. Table 4 summarizes the results of this regression analysis. This model allowed us to test both **H1** and **H2**, but, for clarity, we discuss the results separately for the individual differences measures (**H1**) and those for the effects of note-taking (**H2a** & **H2b**).

Table 4. *Effects of note-taking condition and individual differences on mind wandering*

Predictor	**β**	***SE***	***t***	***p***
H1: Individual Differences Variables				
Pretest	**-0.16**	0.07	-2.21	**.029**
Number of Math Courses	-0.10	0.07	-1.33	.184
Math Interest	**-0.20**	0.08	-2.52	**.013**
Learning Confidence	0.06	0.07	0.79	.430
Incremental Beliefs	-0.06	0.08	-0.81	.419
Media Multitasking Habits	**0.18**	0.07	2.53	**.012**
Note-taking Habits	-0.11	0.07	-1.52	.131
WMC	-0.09	0.07	-1.22	.226
H2: Effects of Note-Taking				
Condition (Note-Taking)	-0.04	0.14	-0.31	.756
Condition × Pretest	**0.32**	0.15	2.21	**.029**
Condition × Number of Math Courses	**0.30**	0.15	2.01	**.046**
Condition × Math Interest	0.06	0.16	0.40	.693
Condition × Learning Confidence	-0.20	0.15	-1.33	.187
Condition × Incremental Beliefs	-0.16	0.15	-1.05	.296
Condition × Media Multitasking Habits	-0.02	0.15	-0.17	.865
Condition × Note-taking Habits	0.17	0.14	1.15	.251
Condition × WMC	-0.16	0.15	-1.11	.268

Note. WMC = working memory capacity. Total R^2 for the model = .24, $F(17, 164) = 3.07$, $p < .001$. Significant coefficients and the corresponding *p* values are printed in bold.

H1: Unique predictors of mind wandering rate. The top half of Table 4 summarizes the regression results for the individual-differences variables. Although, as was shown in Table 3, four measures were significantly correlated with TUT rate, the regression results showed that only three of them uniquely predicted unique variance in TUT rate, beyond the effects of other variables included in the model. Specifically, higher mind-wandering rates were uniquely predicted by lower pretest scores, β = –.16, $t(164) = –2.21$, $p = .029$, lower math interest, β = –.20, $t(164) = –2.52$, $p = .013$, and higher reports of classroom media multitasking, β = .18, $t(164) = 2.53$, $p = .012$. Although significant by itself, $r(180) = –.24$, $p = .001$, incremental beliefs about math ability did not account for significant mind-wandering variance when all the variables were considered simultaneously.

These results are generally consistent with **H1**, insofar as prior math/stats interest (**H1a**), prior knowledge, incremental beliefs about math, and in-class media multitasking habits (**H1b**) all significantly correlated with TUT rate, and all but incremental beliefs accounted for unique variance in mind wandering during the lecture. However, inconsistent with **H1a**, WMC's predicted negative correlation ($r = -.12$) with TUT rate was not significant.

H2a and H2b: Main and moderating effects of note-taking on mind wandering. The bottom half of Table 4 summarizes the regression results involving the effects of the note-taking manipulation ("Condition"). Inconsistent with **H2a** (i.e., that note-taking should help scaffold attention and reduce TUT rates), note-taking condition did not predict mind wandering rate at the overall group level (i.e., no significant main effect of note-taking). In line with **H2b**, however, the note-taking manipulation significantly interacted with two of the individual-differences variables, both of which we consider to reflect background knowledge. Specifically, note-taking interacted with pretest scores, $\beta = .32$, $t(164) = 2.21$, $p = .029$, and with number of math courses taken in high school, $\beta = .30$, $t(164) = 2.01$, $p = .046$. These interactions are graphically illustrated in Figure 2.

Not surprisingly, lower pretest scores (panel a) and the fewer number of math courses taken (panel b), respectively, were associated with greater mind-wandering frequency. But, for both of these variables, the negative effect of those individual differences dimensions (reflected in the slopes of the regression lines) was significantly weaker for subjects who had an opportunity to take notes compared to those who did not, as hypothesized in **H2b**. Indeed, note-taking appears to have eliminated the effect of math courses on TUTs. These moderating influences of note-taking suggest that note-taking helped students with little prior background knowledge stay focused.

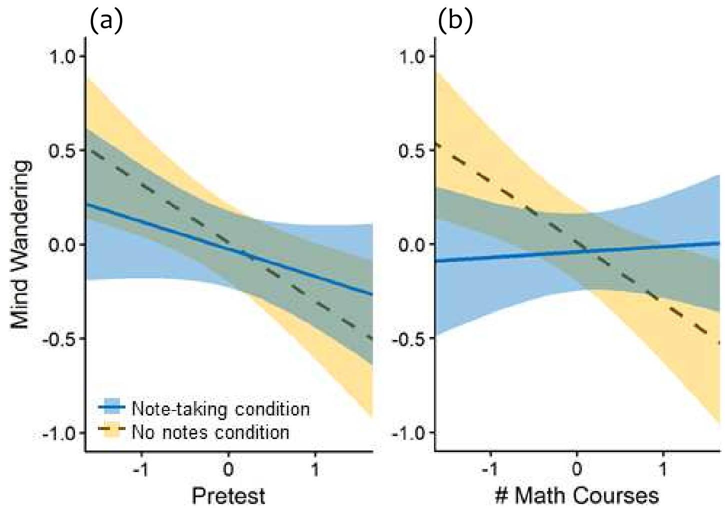

Figure 2. Moderating effects of note-taking on mind wandering for two knowledge-related individual differences variables: (a) pretest and (b) the number of math courses taken in high school. The y-axis represents the standardized (z-scored) mind-wandering rate, operationalized as proportion of time students reported having task-unrelated thoughts (TUTs). Pretest, plotted on the x-axis for panel (a), is the participant's z-scored number of correctly answered questions on the pretest, whereas # Math Courses in panel (b) is the z-scored number of math-related courses the participants reported having taken. Shaded regions represent 95% confidence intervals around the regression lines.

H2c: Effects of note-taking quality/quantity. As stated in the introduction, the main effect of note-taking (**H2a**) may not be significant if not everyone in the note-taking condition took good notes throughout the lecture. Thus, although this is a less powerful test of the potential benefits of note-taking, we hypothesized a correlation such that students who took better (or more) notes should show less mind wandering, better learning, and stronger situational interest than those who took poorer (or fewer) notes (**H2c**).

To test this hypothesis, we analyzed the data from the 94 subjects in the note-taking condition, despite the substantial reduction in power that came from splitting the sample roughly in half. Specifically, as described earlier, we coded the quality/quantity of the notes taken by each subject and used the resulting scores for analyses (*M* = 22.9, out of 31 points; *SD* =

6.8; Range = 5–31). Consistent with **H2c**, the rated note quality/quantity measure was correlated substantially with mind-wandering frequency, $r(92) = -.53$, $p < .001$. As illustrated in Figure 3, those subjects who took better notes indeed experienced fewer TUTs during the video lecture.[7]

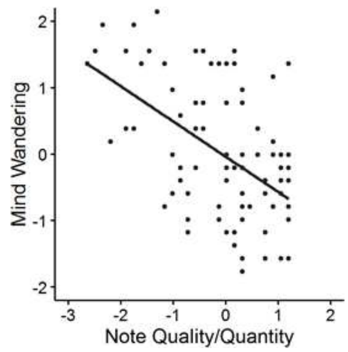

Figure 3. The scatter plot illustrating the bivariate correlation between mind wandering (the rate of task-unrelated thoughts [TUTs]) and note quality/quantity among the subjects in the note-taking treatment condition (n = 94). The y-axis represents the z-scored proportion of time students reported having TUTs during the video lecture. The Note Quality/Quantity variable plotted on the x-axis represents the standardized (z-scored) ratings of the combined quality and quantity of participants' notes.

Table 5. *Effects of note quality/quantity and individual differences on mind wandering*

Predictor	β	*SE*	*t*	*p*
Note Quality/Quantity	**-0.49**	0.10	-4.95	**<.001**
Pretest	0.04	0.09	0.39	.697
Number of Math Courses	0.06	0.09	0.63	.533
Math Interest	-0.14	0.10	-1.34	.185
Learning Confidence	0.02	0.09	0.21	.838
Incremental Beliefs	-0.17	0.10	-1.73	.087
Media Multitasking Habits	0.10	0.09	1.21	.231
Note-Taking Habits	0.03	0.09	0.30	.763
WMC	-0.14	0.09	-1.48	.143
Note Quality/Quantity × Pretest	0.10	0.12	0.87	.385
Note Quality/Quantity × Number of Math Courses	-0.10	0.09	-1.08	.284
Note Quality/Quantity × Math Interest	0.14	0.11	1.30	.196
Note Quality/Quantity × Learning Confidence	0.06	0.09	0.61	.545
Note Quality/Quantity × Incremental Beliefs	0.10	0.11	0.94	.350
Note Quality/Quantity × Media Multitasking Habits	0.17	0.10	1.81	.075
Note Quality/Quantity × Note-Taking Habits	-0.12	0.09	-1.36	.177
Note Quality/Quantity × WMC	-0.05	0.12	-0.44	.660

Note. WMC = working memory capacity. Total R^2 for the model = .48, $F(17, 76) = 4.13$, $p < .001$. Significant coefficients and the corresponding *p*-values are printed in bold.

[7] Note quality/quantity also correlated significantly with posttest scores, $r(92) = .32$, $p = .002$, and situational interest, $r(92) = .38$, $p < .001$, but not with any of the individual differences measures, all *r*s ≤ .18 all *p*s ≥ .080.

We then conducted a regression analysis in which the dependent variable was TUT rate, and the predictor variables were the note quality/quantity variable and all the same individual differences variables summarized in the top half of Table 3. The regression results are reported in Table 5. As hypothesized (**H2c**), note quality/quantity significantly predicted TUT rate during the lecture, $\beta = -.49$, $t(76) = -4.94$, $p < .001$, even with all other correlates of TUT rate included in the model. Indeed, the statistical contribution of note quality/quantity to variance in mind wandering overwhelmed the previously significant effects of pretest, math interest, and media multitasking, such that note quality/quantity was the *only* significant predictor in the model.

H3: Unique Predictive Power of Mind Wandering for Posttest and Situational Interest

Having established the individual-differences and moderating variables that significantly predicted mind-wandering frequency, we then tested **H3** by evaluating, in two separate regression models, the extent to which TUT rate uniquely predicted the two key outcomes (posttest scores and situational interest) above and beyond all the other relevant variables. For these analyses, we ran the same regression models as the one summarized in Table 4, except that (a) the dependent measure was posttest score (averaged across three parts) and situational interest, respectively, and that (b) mind-wandering (TUT) rate was included as an additional individual-differences predictor. The regression results are summarized in Table 6.

Posttest. Of most importance, we found that more mind wandering predicted lower performance in the posttest, above and beyond the contributions of the other predictors in the model, $\beta = -.34$, $t(163) = -5.85$, $p < .001$, including the pretest variable. This result is consistent with **H3**, that mind wandering should predict posttest scores above and beyond the influence of prior knowledge (e.g., pretest) and other variables. As summarized in Table 6 (left panel), additional variables also significantly predicted posttest scores, including pretest performance,

Table 6. *Effects of note-taking condition and individual differences on posttest and situational interest scores with TUT rate as the measure of mind wandering*

	Posttest				Situational Interest			
Predictor	**B**	***SE***	***t***	***p***	**β**	***SE***	***t***	***p***
Mind Wandering Rate	**-0.34**	0.06	-5.85	**<.001**	**-0.47**	0.07	-6.97	**<.001**
Pretest	**0.20**	0.06	3.55	**<.001**	-0.04	0.06	-0.63	.530
Number of Math Courses	-0.07	0.06	-1.26	.211	-0.06	0.06	-0.94	.350
Math Interest	0.12	0.06	1.96	.052	**0.28**	0.07	4.03	**<.001**
Learning Confidence	-0.00	0.06	-0.04	.967	0.01	0.07	0.10	.923
Incremental Beliefs	-0.07	0.06	-1.17	.244	0.01	0.07	0.17	.865
Media Multitasking Habits	-0.00	0.06	-0.07	.943	-0.04	0.06	-0.67	.502
Note-Taking Habits	0.02	0.05	0.30	.768	0.03	0.06	0.47	.639
WMC	0.06	0.06	1.05	.295	-0.10	0.06	-1.53	.128
Condition (Note-Taking)	0.17	0.11	1.60	.112	0.11	0.12	0.92	.358
Condition × Pretest	**0.22**	0.11	1.99	**.048**	-0.05	0.13	-0.39	.697
Condition × Number of Math Courses	-0.03	0.11	-0.29	.769	-0.05	0.13	-0.41	.682
Condition × Math Interest	0.18	0.12	1.55	.123	-0.26	0.14	-1.92	.057
Condition × Learning Confidence	0.03	0.11	0.26	.795	0.26	0.13	1.96	.052
Condition × Incremental Beliefs	0.08	0.12	0.70	.484	0.14	0.13	1.01	.314
Condition × Media Multitasking Habits	0.14	0.11	1.28	.203	-0.02	0.13	-0.17	.865
Condition × Note-Taking Habits	0.00	0.11	0.04	.970	0.10	0.13	0.78	.439
Condition × WMC	0.21	0.11	1.87	.064	0.08	0.13	0.65	.514

Note. WMC = working memory capacity. Posttest model: total $R^2 = .42$, $F(18, 163) = 6.46$, $p < .001$. Situational interest model: total $R^2 = .43$, $F(18, 163) = 6.77$, $p < .001$. Significant coefficients and the corresponding *p*-values are printed in bold.

β = .20, $t(163) = 3.55$, $p < .001$. We again found no main effect of experimental condition, β = .17, $t(163) = 1.60$, $p = .112$, but the note-taking condition interacted significantly with pretest scores, β = .22, $t(163) = 1.99$, $p = .048$. As illustrated in Figure 4, note-taking seems to have enhanced the effects of pretest scores on student learning (posttest performance), especially among those who had higher pretest scores.

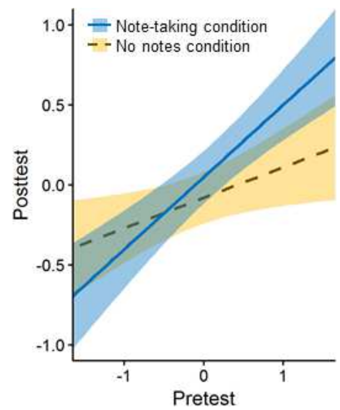

Figure 4. Moderating effect of pretest performance by the note-taking on posttest performance. The y-axis represents the z-scored average of participant's scores on parts 1, 2 and 3 of the posttest. The pretest variable plotted on the x-axis is the participant's z-scored number of correctly answered questions on the pretest. Shaded regions represent 95% confidence intervals around the regression lines.

Situational interest. Similarly, as indicated in Table 6 (right panel), increased reports of in-lecture mind wandering were significantly associated with lower ratings of situational interest in the lecture content, β = –.47, $t(163) = -6.97$, $p < .001$, above and beyond the contributions of the other predictors in the model, including prior interest in math and statistics. Again, this result is consistent with **H3** and suggests that in-lecture TUT rate may be associated with the development of interest during the lecture and not simply reflect the influence of prior interest. Of all other measures in Table 6, only prior math/statistics interest predicted unique variance in situational interest, β = .28, $t(163) = 4.03$, $p < .001$.

H4: Mind Wandering as a Mediator

Data analysis plan. The regression results reported above—namely, that several variables predicted in-lecture TUT rates, which in turn uniquely predicted both of the outcome measures—are consistent with a mediating role of mind wandering (**H4**). We thus conducted two formal mediation analyses, one for each outcome variable, using the *MEDIATE* macro for SPSS (Hayes & Preacher, 2014). This procedure makes it possible to examine multiple predictors by decomposing their "total" effects into "direct" and "indirect" effects. The direct effects are based on the regression models presented for **H3** (Table 6). The indirect effects (with TUT rate as the hypothesized mediator) were considered significant if the 95% bias-corrected bootstrap CIs did not include zero. The mediation results are graphically illustrated in Figure 5.

Posttest. Figure 5 summarizes all significant direct and indirect effects on posttest scores (panel a). Consistent with **H4**, we found significant (positive) indirect effects of pretest scores, indirect = .06, SE = .03, 95% CI [.007, .116], and prior math interest, indirect = .07, SE = .03, 95% CI [.015, .135]. Higher pretest scores were associated with higher posttest scores due, in part, to lower TUT rates, and greater prior interest affected learning indirectly, via its link to decreased TUTs. In addition, classroom multitasking behavior showed a significant (negative) indirect effect on learning through TUTs, indirect = –.06, *SE* = .03, 95% CI [–.120, –.018]: frequent classroom multitasking predicted decreased learning through its positive association with mind wandering. Moreover, the two significant interaction terms in predicting the posttest scores (Table 6) also showed significant indirect effects: the note-taking condition × pretest interaction, indirect = –.11, SE = .05, 95% CI [–.229, –.012], and the note-taking × math courses interaction, indirect = –.10, SE = .05, 95% CI [–.205, –.013]. No other variables in the model (i.e., the variables listed in Tables 4 and 6) showed significant indirect or direct effects.

Taken together, consistent with **H4**, the predictive power of all of these five variables for the posttest scores was significantly mediated by TUT rate. As shown in Figure 5a, however, two

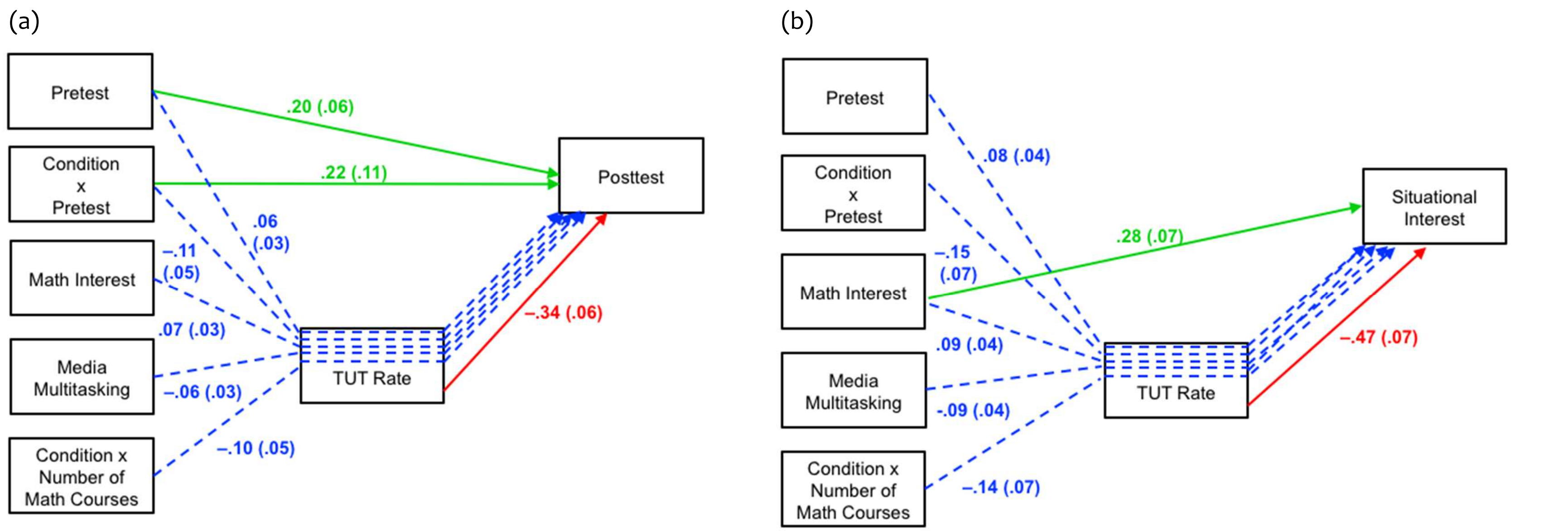

Figure 5. Schematic summary of the results of the mediation analyses, testing the hypothesized mediating role of mind wandering (the rate of task-unrelated thoughts [TUTs]), for (a) the posttest outcome variable and (b) the situational interest outcome variable. Significant indirect effects (through TUT rate) are indicated by dotted blue lines, and direct effects are indicated in solid lines, with positive direct effects indicated by green and negative effects printed in red (standard errors or SEs in parentheses). These mediation analyses were conducted with the MEDIATE macro for the SPSS statistical package and included all the predictor variables summarized in Table 6. Only those predictors that demonstrated significant indirect or direct effects are illustrated in these diagrams.

of the five variables—pretest score and the pretest score x note-taking interaction—also demonstrated significant direct effects, meaning that the predictive power of the pretest scores (and its interaction with the note-taking manipulation) went above and beyond its impact on mind wandering. This finding is not particularly surprising, given that pretest scores were expected to be the best predictor of posttest scores.

Situational Interest. We also tested whether any predictors of situational interest or mind wandering demonstrated indirect effects on situational interest via their influence on TUTs. As illustrated in Figure 5b, all five predictors showed significant indirect effects. Both higher pretest scores, indirect = .08, SE = .04, 95% CI [.006, .172], and greater math interest, indirect = .09, SE = .04, 95% CI [.017, .190], predicted greater situational interest due in part to their association with lower TUT rates, although, as shown in Figure 5b, this latter variable had a significant direct effect as well. Moreover, more classroom multitasking predicted decreased interest via higher TUT rates, indirect = –.09, SE = .04, 95% CI [–.167, –.022]. Negative indirect effects were also seen for the interactions: condition × pretest, indirect = –.15, SE = .07, 95% CI [–.304, –.012], and condition × math courses, indirect = –.14, SE = .07, 95% CI [–.295, –.011]. No other variables in the analyses demonstrated significant direct or indirect effects.

These results are consistent with **H4** and demonstrate that the predictive power of the five variables summarized in Figure 5b was mediated by mind-wandering frequency during the lecture. Only prior math interest, which correlated substantially (r = .39) with situational interest (Table 3), demonstrated an additional direct effect that was independent of TUT rate.

H5: Potential Beneficial Effects of Topic-Related Mind Wandering

Thought probes presented during the video lecture allowed subjects to report not only about unambiguously on- and off-task thoughts, but also about two types of thoughts that were related to the lecture material but not to what was being discussed in the video at that moment. That is, subjects could have (a) *lecture-related thoughts*, which were about statistics, or something from the video, but not focused on the here-and-now of the ongoing lecture, or (b) *comprehension-related thoughts,* which were about how well or how poorly subjects felt they understood the material (see Figure 1). The inclusion of these two off-task—but topic-related—probe options allowed us to test the hypothesis (**H5**) that students who spend considerable time thinking about meaningfully related material (having more lecture-related thoughts), but not about their own learning or comprehension (having more comprehension-related thoughts), might actually perform better, despite being occasionally tuned out from the ongoing lecture.

Table 7. *Zero-order correlations between the frequencies of two types of off-task thoughts (lecture-related and comprehension-related) and the outcome variables*

Measure	***TUT Rate***	***Posttest***	**Situational Interest**
Lecture-Related Thoughts	**−.41**	**.26**	**.15**
Comprehension-Related Thoughts	**−.24**	.01	**.15**

Note. Degrees of freedom (*df*s) for the correlations = *180.* TUT = task-unrelated thoughts.

Zero-order correlations. The main results of our planned correlational analysis are summarized in Table 7. Note that, because having more lecture-related or comprehension-related thoughts necessarily meant experiencing fewer off-task thoughts of other types, it is not surprising that the overall rate of mind wandering about unrelated topics during the lecture (TUT rate) was negatively correlated with both lecture-related thoughts, $r(180) = -.41$, $p < .001$, and comprehension-related thoughts, $r(180) = -.24$, $p = .001$.[8]

Of most importance, however, subjects who had more lecture-related thoughts also

[8] This statistical dependence between TUTs and lecture-/comprehension-related thoughts—which reflected mutually exclusive response choices at each thought probe—led us to analyze these thought types separately in the exploratory follow-up regression analyses reported below and summarized in Table 8.

scored better on the posttest, $r(180) = .26$, $p < .001$, and found the lecture more interesting, $r(180) = .15$, $p = .046$, than did those who reported fewer lecture-related thoughts. In contrast, although comprehension-related thoughts correlated positively with situational interest $r(180) = .15$, $p = .043$, it was not correlated with posttest scores, $r(180) = .01$, $p = .923$. A Fisher's *r* to *z'* transformation test for nonindependent samples indicated that the correlation with the posttest scores was significantly higher for lecture-related thoughts (.26) than for comprehension-related thoughts (.01), $t = 2.42$, $p = 0.02$. Consistent with **H5**, these results suggest that certain types of off-task thoughts (i.e., lecture-related thoughts) are beneficial for—or at least are associated with—the enhancement of learning and situational interest.

Exploratory follow-up analyses. To further examine the potential consequences of lecture-related mind wandering on learning and situational interest, we conducted parallel regression analyses to those we reported above for TUTs. The results of these exploratory analyses are summarized in Table 8 for the models predicting lecture-related mind wandering (left panel) and posttest scores (right panel).

As shown in the left panel of Table 8, lecture-related mind wandering was significantly predicted by note-taking condition, $\beta = -.47$, SE = 0.14, $t(164) = -3.36$, $p = .001$, but, interesting to note, *more* lecture-related thought was observed for subjects in the *no-notes* condition than for those in the note-taking condition. Perhaps not having to take notes gave those no-note subjects more opportunity for such topic-related reflections. In addition, lecture-related mind wandering was significantly predicted by pretest scores (with more lecture-related thought for students who did better on the pretest), $\beta = .18$, SE = 0.07, $t(164) = 2.46$, $p = .015$, and by the interaction between note-taking condition and note-taking habits (such that note-taking reduced lecture-related mind wandering for those with better note-taking habits, whereas

Table 8. *Results of exploratory follow-up regression analyses on direct effects of note-taking condition and individual differences on posttest and situational interest scores with lecture-related thoughts as the measure of mind wandering*

	Lecture-Related Mind Wandering				Posttest			
Predictor	**β**	***SE***	***t***	***p***	**β**	***SE***	***t***	***p***
Lecture-Related Mind Wandering	–	–	–	–	**0.13**	0.06	1.99	**.049**
Pretest	**0.18**	0.07	2.46	**.015**	**0.23**	0.06	3.79	**<.001**
Number of Math Courses	0.11	0.07	1.45	.139	-0.05	0.06	-0.83	.409
Math Interest	0.06	0.07	0.76	.449	**0.18**	0.06	2.77	**.006**
Learning Confidence	0.09	0.08	1.19	.235	-0.03	0.06	-0.55	.581
Incremental Beliefs	0.11	0.08	1.48	.141	-0.06	0.06	-0.96	.338
Media Multitasking Habits	-0.09	0.07	-1.23	.221	-0.06	0.06	-0.94	.351
Note-Taking Habits	-0.02	0.07	-0.27	.787	0.06	0.06	0.95	.342
WMC	0.03	0.07	0.40	.688	0.08	0.06	1.42	.158
Condition (Note-Taking)	**-0.47**	0.14	-3.36	**<.001**	**0.24**	0.12	2.05	**.042**
Condition × Pretest	0.04	0.15	0.28	.777	0.11	0.12	0.89	.376
Condition × Number of Math Courses	-0.14	0.15	-0.92	.359	-0.12	0.12	-0.97	.332
Condition × Math Interest	-0.05	0.15	-0.33	.743	0.17	0.13	1.31	.191
Condition × Learning Confidence	0.09	0.15	0.57	.573	0.09	0.12	0.71	.479
Condition × Incremental Beliefs	0.03	0.15	0.19	.851	0.13	0.13	1.06	.291
Condition × Media Multitasking Habits	0.02	0.14	0.12	.902	0.15	0.12	1.23	.221
Condition × Note-Taking Habits	**-0.31**	0.14	-2.18	**.031**	-0.01	0.12	-0.11	.912
Condition × WMC	0.25	0.15	1.69	.093	0.23	0.12	1.91	.057

Note. WMC = working memory capacity. Lecture-related mind wandering model: total $R^2 = .24$, $F(17, 164) = 3.13$, $p < .001$. Posttest model: total $R^2 = .31$, $F(18, 163) = 4.08$, $p < .001$. Significant coefficients and the corresponding *p*-values are printed in bold.

not taking notes increased lecture-related mind wandering for those with better note-taking habits), $\beta = -.31$, SE = .14, $t(164) = -2.18$, $p = .031$.

As shown in the right panel of Table 8, some predictor variables—pretest scores, math interest, and note-taking—significantly predicted students' posttest performance. Most critically, however, even with all the other predictor variables in the models, lecture-related mind wandering still significantly positively predicted posttest scores, $\beta = .13$, SE = 0.06, $t(163) = 1.99$, $p = .049$. A parallel regression analysis conducted for the situational interest outcome (not summarized in the table for simplicity), however, did not uniquely predict situational interest, $\beta = .05$, SE = .08, $t(163) = 0.68$, $p = .501$.

Finally, a follow-up mediation analysis (again not graphically illustrated for simplicity), conducted only for the posttest measure, indicated that all the predictors of lecture-related mind-wandering also had significant indirect effects (via lecture-related mind wandering) on posttest scores: for pretest scores, indirect = .023, SE = .016, 95% CI [.001, .064], for note-taking condition, indirect = −.06, SE = .03, 95% CI [−.149, −.006], and for the note-taking condition × note-taking habits interaction, indirect = −.04, SE = .03, 95% CI [−.130, −.001].

Consistent with **H5**, then, these results suggest that lecture-related mind wandering can tell us something unique about students' learning from an on-line lecture. Moreover, lecture-related mind wandering seems to modestly mediate other predictors' influence on learning.

Testing the Generalizability of the Main Results Across Two Research Sites

In the regression and mediation results reported above, we did not assess potential differences between the two research sites because our primary motivation in conducting this multisite study was to ensure sufficient variability in all the outcome and individual-differences measures we collected. We also assumed that our main findings would be robust to any particular site-specific effects. Here, we briefly report the results of some supplementary analyses we conducted to test this generalizability assumption.

Descriptive statistics and zero-order correlations. Appendix A reports the mean differences between research sites for each outcome and predictor variable, collapsed, for simplicity, across note-taking condition.[9] Students at the two sites differed substantially in all outcome measures: Differences in TUT rate ($d = -.77$) and posttest scores ($d = 1.00$) were substantial, but the difference in situational interest was more modest ($d = .41$). The site differences were generally absent for the individual-differences predictor variables, with only two showing modest effects: pretest ($d = .35$) and WMC ($d = .33$).

Appendix B summarizes the correlations among measures for each site (with Site A correlations below the diagonal and Site B correlations above). There were some notable differences in the patterns of correlations (e.g., the incremental beliefs variable significantly predicted mind wandering and situational interest at Site B, but not at Site A), suggesting the presence of some site-specific effects. It is important to emphasize, however, that some key correlations were significant for both sites (e.g., mind-wandering rate was substantially correlated with both posttest scores and situational interest).

Regression models involving Site. Of primary importance—to examine the extent to which the key results we reported above were robust to any site-specific effects—we conducted a set of regression analyses analogous to those reported in Table 4 (for mind wandering) and Table 6 (for posttest and situational interest), but including Site as an additional fixed-effect variable (coded as –0.5 for Site B and +0.5 for Site A) and fully crossing Site with both note-

[9] As summarized in Appendix C, no Site × Condition interactions were significant for any of the outcome measures, hence justifying the reporting of the site results collapsed across Condition.

taking condition and all of the individual-differences predictors. Specifically, the model for each outcome variable included all the terms included in the models shown in Tables 4 and 6 and all of the two-way and three-way interactions involving Site. If the key results reported above could be similarly observed across the two sites, then the significant predictors in the models summarized in Tables 4 and 6 should still remain significant, even when the Site differences are explicitly modeled, and despite some site-specific effects appearing as higher-order (two-way or three-way) interactions involving Site. As summarized in Appendix C, this is what we found.

As expected in light of significant site differences in the outcome variables (Appendix A) and in the bivariate correlations (Appendix B), regression models that included Site (and its interactions) accounted for more total variance in each outcome measure compared to the primary models summarized in Tables 4 and 6: an increase of .14 for mind wandering (from .24 to .38), .13 for posttest scores (from .42 to .55), and .06 for situational interest (from .43 to .49). Moreover, Site was a strong predictor of mind wandering ($\beta = -.54$) and posttest scores ($\beta = .50$), but not of situational interest ($\beta = .01$).

Despite such substantial effects of Site, however, virtually all of the predictors that were significant in our primary analyses summarized in Tables 4 and 6—indicated by boxes surrounding the statistics in Appendix C—remained significant. The only two exceptions involved small changes to *p* values from one side of the alpha level to the other: (a) math interest significantly predicted mind wandering before ($p = .013$; Table 4), but not with Site in the model ($p = .063$; Appendix C, left panel); and, similarly, (b) the Condition × Pretest interaction was significant before ($p = .048$; Table 6, left panel), but not with Site in the model ($p = .075$; Appendix C, middle panel). Moreover, the regression results indicated very few site-specific effects in the form of significant two- and three-way interactions involving Site (e.g., Site ×

Incremental Beliefs for mind wandering; Site × Condition × Math Interest for posttest scores; Site x Pretest for situational interest), despite the large number of interactions tested (i.e., 6 significant interactions involving Site out of 51 tested).[10]

On the basis of these analyses, we can conclude that the key results reported previously for our main analyses (Tables 4 and 6) were robust and generalizable across sites. Because most previous educationally relevant and/or individual-differences studies of mind wandering have involved only a single site, it is difficult to gauge the extent which their findings are generalizable to other samples. Therefore, the current demonstration of some generalizability of results across two quite different samples—despite some clear site-specific effects—gives some credence to the likelihood of replicability of our key results in new samples and thereby adds substantially to the novelty and information value of the present study.

Discussion

College students in the U.S. increasingly enroll in online coursework, with about a third taking at least one online class (e.g., Allen & Seaman, 2013); many are therefore learning from video-based materials. Although a small literature has examined the costs of mind wandering during recorded lectures (as well as live lectures and reading materials), few studies have examined individual differences in the predictors and consequences of mind wandering. None, moreover, have rigorously assessed whether such individual differences might moderate the

[10] We also conducted regression analyses separately for each site after standardizing (z-scoring) all the relevant variables for each site. These results are available in the supplementary materials (as is the case with Appendix C, the predictors that were significant in our main analyses are indicated by boxes in Tables S1–S3). Likely due in part to the substantial reduction in statistical power resulting from splitting the sample roughly in half, the results of these site-by-site analyses are less clear-cut, especially for the mind wandering models (Table S1), where no single predictor remained significant (with the exception of the significant effect of media multitasking observed at Site B). Of importance, however, even these noisier site-by-site analyses demonstrated that, although some clear across-site differences are present for some predictor variables, many of the predictors that were significant in the main analyses (Table 6) were still significant (or close to being significant) at both sites for the posttest and situational interest models (see Tables S2 and S3, respectively).

beneficial effects of particular educationally relevant activities (such as note-taking) for reducing mind wandering (i.e., individual-by-treatment interactions). And no prior studies that we know of have systematically investigated how mind-wandering experiences might affect the interest that people derive from their ongoing activities (i.e., situational interest): In the classroom, the disruption of processing and engagement that results from frequent mind wandering might not only impair students' learning from what they see, hear, and read, but it may also lessen their enthusiasm for future study of that topic.

Our laboratory study, using a diverse sample, took a combined experimental–correlational approach to extending prior research on mind wandering in educationally relevant contexts. We found a robust average rate of mind wandering during a video lecture on introductory statistics (*M* = 46%), and considerable individual-differences variation around that average rate (*SD* = 26%). Of most importance, we tested five *a priori* hypotheses about in-lecture mind wandering and thus organize our subsequent discussion around them.

H1: Unique Predictors of Mind-Wandering Rate

Prior educationally relevant studies of individual differences in TUT rates have found that they correlate with WMC, topic interest, or both (Hollis & Was, 2016; Lindquist & McLean, 2011; McVay & Kane, 2012b; Unsworth & McMillan, 2013); when WMC and interest are examined together, they each predict mind wandering independently. The challenge in interpreting the findings about interest, however, is that it has often been assessed only *after* learning (or as a combined pre/post measure). We thus cannot know whether students' prior interest contributes to their subsequent TUT rates during learning, or whether the experiences of mind wandering or learning influence students' assessments of their topic interest. In contrast, the present study measured interest in mathematics and statistics *before* students watched the video lecture.

Moreover, previous research has considered few other potential predictors of mind wandering. Prior knowledge has shown mixed evidence (Unsworth & McMillan, 2013; Wammes et al., 2016b), but those studies relied on only self-report measures; we therefore assessed prior knowledge via a pretest of the material, in addition to self-reports of prior coursework. We also measured students' beliefs that math ability is malleable, their confidence in their own ability to learn the material, and their habitual engagement in media multitasking during classes.

Consistent with prior studies, we found no correlation between WMC and prior math interest and a significant negative correlation between prior interest and TUT rate. In contrast to **H1a** and also to some past work, however, we found a nonsignificant negative correlation between WMC and TUTs during the lecture ($r = -.12$). Our regression analyses of TUT rate, which pitted WMC and prior interest against one another (and against the other individual-differences variables), found that math interest, but not WMC, uniquely predicted TUTs (Table 4).

Why did we find no statistical evidence for the predictive power of WMC? Although many studies find that higher-WMC subjects mind-wander less than do lower-WMC subjects, these associations are weak, with bivariate correlations in the −.10 to −.15 range. Indeed, a recent meta-analysis (Randall, Oswald, & Beier, 2014) estimated the WMC–TUT correlation to be −.12 (95% CI [−.18, −.06]); if this estimate is correct, then our study was underpowered to detect a statistically significant WMC–TUT rate association (note that our effect size estimate matched the meta-analytic one). Future research should either use multiple WMC and TUT assessments to allow latent-variable approaches, where WMC–TUT correlations are stronger (Hollis & Was, 2016; McVay & Kane, 2012b; Unsworth & McMillan, 2013), or should focus on attention-control tasks as cognitive predictors, which also correlate more strongly with TUT rates (Kane et al., 2016; McVay

& Kane, 2012a; Unsworth & McMillan, 2014).

Whereas pretest scores, classroom media multitasking habits, and incremental beliefs about math ability also correlated with TUT rate, only pretest and media multitasking (along with prior math interest) predicted unique variance in TUTs (**H1b**). Although the negative association between pretest scores and TUT rates contradicts the null correlation that Unsworth and McMillan (2013) found between self-reported knowledge and TUTs during reading, it jibes with the Wammes et al. (2016b) report of a significant negative correlation ($r = -.19$) between self-reported prior knowledge and unintentional TUTs during classroom learning (prior knowledge, however, did not correlate with intentional mind wandering; $r = .02$). These findings, including ours, indicate that students who enter the learning context with a stronger background in the topic can better sustain attentional focus on the lecture, perhaps because they can better link new, incoming information to extant knowledge structures to create better mental models (e.g., Smallwood et al., 2008).

Perhaps our most novel individual-differences finding is that students' endorsement of media multitasking in classroom contexts predicted TUTs during learning (**H1b**), even though the lab context did not allow for this behavior, nor did it present cues because subjects had to silence and store their phones. This finding is broadly consistent with a growing body of research, using more general media multitasking questionnaires, that shows that higher endorsement of media multitasking habits is associated with lower performance of computerized attention tasks (e.g., Cain et al., 2016; Moisala et al., 2016; Ophir et al., 2009). As well, students who endorse more frequent media multitasking—particularly in academic contexts—show poorer academic performance (e.g., Fox, Rosen, & Crawford, 2009; Junco & Cotton, 2012; Kraushaar & Novak, 2010). The current study adds new evidence based on TUT rate to this growing literature on the negative effect of media multitasking habits. At this point, however, the precise relationship between classroom media multitasking and in-lecture TUT rate is not clear, given the correlational nature of the study. Further research is needed to test whether this relationship is causal in nature.

H2: The Effects of Note-Taking on Mind Wandering

Contrary to **H2a**, we did not find the predicted benefits of note-taking at the group level: Our experimental note-taking manipulation did not have significant overall (main) effects on TUT rate, posttest scores, or situational interest. The regression results (summarized in Tables 4 and 5), however, suggest two important ways in which the note-taking manipulation had some significant impacts on subjects' in-lecture mind wandering.

First, consistent with **H2b**, we found two individual-by-treatment interactions, both involving aspects of prior knowledge (Table 4): Note-taking reduced the cost of low pretest scores on TUT rate, where lower pretest scores tended to predict more mind wandering (Figure 2a) and minimized the cost of having taken fewer prior math courses on TUT rate, where fewer courses also tended to predict more mind wandering (Figure 2b). Although we did not predict that these two prior-knowledge-related variables (versus others in the analysis) would significantly interact with the note-taking manipulation, the specific patterns of interactions are in line with our *a priori* hypothesis (**H2b**): An opportunity to take notes should reduce the impacts of some individual-differences variables on in-lecture mind wandering, as indicated by the shallower slopes in Figures 2a and 2b for the note-taking than for the no-notes condition. Moreover, these findings are broadly consistent with prior studies showing that note-taking benefits may depend on student characteristics (Einstein, Morris, & Smith, 1985; Peper & Mayer, 1978). However, given that we did not specifically predict that note-taking's benefits would be focused on students with less prior knowledge, it is important to replicate these findings.

Second, consistent with **H2c**, our analysis of the quality/quantity of students' notes provided further evidence for the relevance of note-taking. We obviously could not force students in the note-taking condition to take effective notes, and so we had to consider how normal variation in note-taking might predict TUT rates. In fact, as shown in Figure 3, students who took more complete and effective notes also mind wandered less than did those who took less complete and effective notes. This finding is consistent with results from Lindquist and McLean (2011), who found that *self-reported* quantity of notetaking in prior course meetings correlated negatively with TUT rate during a live lecture (Kendall's $\tau = -.16$), but also provides more robust, objective evidence. Moreover, we found that variation in note quality/quantity predicted TUT rates in our regression analysis that included all of our other predictor variables in the model (Table 5). Thus, note-taking was not simply a stand-in for the other individual-differences factors that we measured.

With that said, we must be cautious in entertaining causal claims regarding the effect of note quality/quantity on TUT rate because we did not manipulate it experimentally. It is possible, for example, that having a stronger attentional focus allowed students to take better notes on the lecture, rather than effective note-taking helping students maintain their focus. In addition, although note quality/quantity predicted mind wandering above and beyond the other predictors measured in our study, it is possible that another third variable (such as motivation) may have caused variation in note quality/quantity that resulted in its association with TUT rate.

H3 and H4: Unique Predictive Power of Mind Wandering and Its Mediating Role

Not only did lecture TUT rate correlate strongly with both posttest scores and situational interest derived from the video ($rs \approx -.50$), but, consistent with **H3**, it also incrementally predicted both outcomes beyond the other predictors in our regression models. Indeed, lecture TUT rate explained significant posttest variance above and beyond that accounted for by pretest scores, and it explained significant situational interest variance above and beyond that accounted for by prior math/statistics interest. These findings point to the robustness of mind-wandering propensity as a predictor of academic outcomes.

Consistent with **H4**, in-lecture mind wandering also significantly mediated *all* of the individual-differences effects in our data. It is not surprising that students' pretest scores and the pretest × note-taking interaction each had additional direct effects on posttest scores, or that students' prior math interest had a direct effect on situation interest, given the similarity of predictors to outcomes. We find it noteworthy, however, that they also had indirect effects on these outcomes via their association with TUT rate during the lecture (Figures 5a & 5b). That is, one cannot fully explain the beneficial influence of prior knowledge on learning from lectures without considering how prior knowledge might help students maintain their attentional focus on the to-be-learned material; likewise, one cannot fully explain how prior interest translates into triggering situational interest in a lecture on that topic without considering how prior interest helps keep students mentally on task. Moreover, the other individual-differences predictors (including multitasking habits and the math courses × note-taking interaction) had *only* indirect effects on posttest scores (Figure 5a) or situational interest (Figure 5b), via their associations with TUT rate (i.e., no direct effects). In particular, the finding that media multitasking tendencies had only an indirect association with learning maps nicely onto a prior finding based solely (and nonoptimally) on retrospective questionnaire measures, that student reports of texting in class had an indirect association with learning that was fully mediated by self-reported sustained-attention ability (Wei, Wang, & Klausner, 2012).

H5: Potential Beneficial Effects of Topic-Related Mind Wandering

Sometimes students' thoughts during an ongoing lecture are completely off-topic. But thoughts during class can also be topic-related while not focused on the here-and-now: Students sometimes reflect on earlier portions of a lecture or on their prior knowledge while learning something new, and they sometimes have metacognitive thoughts about how well they are understanding what they are seeing and hearing. We consider these lecture- and comprehension-related thoughts to be forms of mind wandering because they reflect at least a partial mental departure from the present moment. At the same time, they also may have quite different antecedents and consequences than do more prototypical examples of off-topic daydreaming or woolgathering. Indeed, comprehension-related and especially lecture-related thoughts might indicate particularly deep engagement in the lecture material.

Consistent with prior classroom studies (Locke & Jensen, 1974; Schoen, 1970), we found that students reported lecture-related thoughts and comprehension-related thoughts at nontrivial rates (*M*s = 15% and 9%, respectively; Figure 1). Also, as in the Frank et al. (2015) study of reading, students' rates of comprehensions-related thoughts told us nothing about their learning, though they did positively correlate with situational interest in the lecture (Table 7).

Of primary interest, our findings support **H5** and are also consistent with those of Jing et al. (2016), who found lecture-related thought rates to positively correlate ($r \approx .45$) with test performance on lecture materials. The correlations of lecture-related TUTs with posttest scores and situational interest were not nearly as strong in our dataset (*r*s = .26 and .15, respectively), probably because the Jing et al. sample was much smaller (n = 36) and thus yielded a less precise (and likely overestimated) effect size. Lecture-related mind wandering, then, may not be as strongly positively associated with learning and situational interest as off-topic mind wandering is negatively associated with them (indeed, in exploratory regression analyses reported above, lecture-related mind wandering significantly predicted only posttest scores and not situational interest). However, the modest positive correlations we found indicate that attentionally engaged learners not only focus their attention on the ongoing material, but they also occasionally "check out" from the here-and-now to think further about the topic, whether about prior portions of the lecture or connections to prior knowledge. Along with findings that future-oriented mind wandering and useful off-task thoughts may improve mood (Franklin et al., 2013; Ruby, Smallwood, Engen, & Singer, 2013), our lecture-related-thought findings thus provisionally suggest that some mind-wandering content can be beneficial (**H5**).

Our data are also unique in speaking to potential contextual and individual predictors of lecture-related mind wandering. As summarized in Table 8 (left panel), lecture-related mind wandering, like TUTs, was predicted by pretest scores, but in the opposite direction: Students who knew more about statistics prior to the lecture reported more lecture-related thoughts during the video than did those who knew less. Moreover, unlike TUTs, these on-topic experiences were affected by our experimental manipulation, with note-taking subjects as a group reporting significantly *less* lecture-related mind wandering than no-notes subjects (Table 8, left panel). An interaction of note-taking condition with note-taking habits further indicated that this dampening effect of note-taking was strongest for students self-reporting better note-taking skills. It seems, then, that having to keep up with the demands of taking good quality notes made it more difficult to also engage in extra elaborative thoughts about the lecture.

Issues of Causality

The central, intuitive claim that follows from our findings is that mind wandering during a lecture (at least TUTs) causally disrupts learning and the development of situational interest.

However, all of the critical data involving mind wandering rate in this study are correlational, and so they are potentially amenable to alternative causal hypotheses. Thus, here we consider some relevant evidence regarding the potential causal influences of mind wandering.

Mind wandering and learning. One alternative to the causal hypothesis is that TUTs actually have no effect on learning, but rather students who experience learning difficulties during a lecture eventually and increasingly tune out from it. The literature on reading and mind wandering has empirically addressed this directional ambiguity in two ways: (a) periodically probing for comprehension in the moment and finding that mind wandering at particular parts of the text predicts poorer comprehension of those same parts, and so mind wandering does not reflect a global comprehension deficit (Schooler et al., 2004; Smallwood et al., 2008); and (b) assessing TUT propensity across both reading and nonreading tasks and finding that this *general* propensity predicts comprehension, and so the mind-wandering assessment cannot be driven by reading ability (McVay & Kane, 2012b). Such findings are consistent with the claim that mind wandering during lectures disrupts learning, but those methods are not available for this study.

Instead, to help disambiguate the learning–TUT association, we can first note that, unlike prior research, the present study rigorously measured prior knowledge and prior topic interest. If variation in TUT rate simply reflected the influence of prior knowledge or topic interest, then it should not have predicted our outcomes beyond these other ostensible causal variables (Table 6, left panel). Second, we can consider TUT rate over the earliest portions of the video lecture, when the content was quite straightforward and students should have had little difficulty understanding the material. To do this, we assessed TUT rate across only the first quarter of the thought probes (first 5 of the 20). Despite this more limited sample of behavior and more restricted range of scores, TUTs indeed significantly correlated with the posttest scores, $r(180)$ =

−.39, $p < .001$, and this correlation was nearly as strong as the overall correlation based on all 20 probes, $r(180) = -.48$, $p < .001$. These findings therefore suggest that it is more likely that TUTs influenced learning than learning difficulties influenced TUTs, although further research will be needed to more completely rule out the alternative causal interpretation.

Mind wandering and situational interest. Causal questions are murkier for the situational interest outcome. We measured situational interest after the video, and mind wandering during the video, and so it is tempting to conclude that mind wandering reduced students' subsequent situational interest. However, situational interest may have had a dynamic, cyclical association with mind wandering, such that ongoing levels of interest influenced TUTs and that these episodes of tuning out the lecture then reduced perceived interest.

At the same time, we do not believe that TUTs or postvideo situational interest were driven *entirely* by students' early assessments of their interest in the video. This is because TUT rate predicted situational interest more strongly than—and statistically above and beyond—the influence of prior interest in mathematics/statistics, *r*s = −.56 and −.39, respectively (Table 3). Moreover, the findings that (a) mind wandering mediated the association between prior math interest and situational interest, and (b) TUTs more strongly predicted situational interest in the video than prior math interest predicted TUTs, *r*s = −.56 vs. −.27, respectively, suggest that TUT vulnerability had at least some causal influence on students' interest in the video lecture.

Despite the suggestion of some causal effect of in-lecture mind-wandering on situational interest, it is also important to take into consideration that students' TUT rates during a lecture were also predicted by prior interest in the topic (Tables 3 and 4). Therefore, our findings—and our broader methodological approach to examining TUTs during learning—are quite compatible with feed-forward-and-feedback models of interest and engagement: Traits, habits, interests and abilities affect momentary motivations to learn, and these then create emotional and cognitive states (including focused or distracted attention), which in turn act as mediators to affect learning, which may then alter those original traits, habits, interests, and abilities (e.g., Hidi, 1990, 1995; Hidi & Reninger, 2006; Schiefele, 1991).

Although this bidirectional (rather than simple causal) interpretation is speculative, it provides a basis for future attempts to develop and test a theoretical account of the effects of mind wandering on situational interest. In this regard, TUT-sampling methods of the sort used in the current study will be a useful research tool in elucidating a possible mechanism by which interest, engagement, and boredom may affect students' learning in the moment (Fulmer et al., 2015; Pekrun, Goetz, Daniels, Stupinsky, & Perry, 2010), because they can provide a more direct measure of attentional focus and, hence, dynamic changes in students' awareness in real time.

Limitations and Future Directions

The primary strengths of the current study were: its consideration of five theoretically driven hypotheses about mind wandering, its large and diverse sample of subjects, the successful demonstration of the generalizability of our primary results across two research sites, its focus on a realistic educational situation (learning from a video lecture), its broad assessment of educationally relevant individual differences, its test for individual-by-treatment interactions, its consideration of several forms of mind wandering, and its prelearning assessment of both topic knowledge and interest. At the same time, we acknowledge several limitations of the present work, in addition to those already considered above.

First, we used a single lecture stimulus, so our results may not generalize to other lectures or topics (although we replicated several prior findings using different learning materials). Second, although the video lecture was as long as a typical live lecture (~50 min), it was longer than most videos from online courses (e.g., Khan Academy), which tend to be segmented into smaller units that can be paused between or within segments; students in typical online courses may therefore mind-wander less, with less learning cost, than our subjects demonstrated. Future work should compare continuous versus segmented versions of video lectures. Third, although most of our individual-differences measures used multiple items to tap their construct of interest, we used only one (necessarily error-prone) instrument to assess each construct and thus may have underestimated some of the associations among them; future work should use multiple measures to allow latent-variable analyses. Fourth, as discussed in the Methods section, because there were no previous studies that included—and combined—similar experimental and individual differences variables within a single study, it was not possible to conduct any formal *a priori* power analysis to determine the sample size for the current study (beyond aiming for an N of 200 that would provide adequate power and a useful effect size estimate for most bivariate correlations of interest). Thus, it is possible that our study lacked statistical power to detect some more subtle effects. Finally, because our subjects participated in the study to fulfill a course requirement, they may have been generally less motivated—and therefore more distractible—than would students trying to learn statistics in a for-credit course.

We therefore encourage investigators to expand mind-wandering research further into authentic classroom contexts, while also rigorously assessing cognitive and noncognitive individual differences, in-the-moment mind wandering experiences, and performance- and engagement-relevant outcomes (e.g., Lindquist & McLean, 2011; Wammes et al., 2016a, 2016b). Both classroom and laboratory research should also continue investigating whether educational practices such as taking comprehensive notes, quizzing (Jing et al., 2016; Szpunar et al., 2013), priming students' core values (Cohen, Garcia, Apfel, & Master, 2006; Miyake et al., 2010) as well as the utility values of the course content (Hulleman, Godes, Hendricks, & Harackiewicz, 2010), helps students better focus their attention on classroom activities and thereby better learn course material.

References

Al-Balushi, S. M., & Al-Harthy, I. S. (2015). Students' mind wandering in macroscopic and submicroscopic textual narrations and its relationship with their reading comprehension. *Chemistry Education and Practice, 16*, 680–688.

Allen, I. E., & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States*. Newburyport, MA: Sloan Consortium.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology, 106*, 20–29.

Ambrose, S., Bridges, M., Lovett, M., DiPietro, M., & Norman, M. (2010). *How learning works: 7 research-based principles for smart teaching.* San Francisco, CA: Jossey Bass.

Benassi, V. A., Overson, C. E., & Hakala, C. M. (2014). *Applying science of learning in education: Infusing psychological science into the curriculum.* Retrieved from the Society for the Teaching of Psychology web site: http://teachpsych.org/ebooks/asle2014/index.php

Brown, P. C., Roediger, H. L., & McDaniel, M. A. (2014). *Make it stick: The science of successful learning*. Cambridge, MA: Harvard University Press.

Bui, D. C., & Myerson, J. (2014). The role of working memory abilities in lecture note-taking. *Learning and Individual Differences, 33*, 12–22.

Bui, D. C., Myerson, J., & Hale, S. (2013). Note-taking with computers. Exploring alternative strategies for improved recall. *Journal of Educational Psychology, 105*, 299–309.

Bunce, D. M., Flens, E. A., & Neiles, K. Y. (2010). How long can students pay attention in class? A study of student attention decline using clickers. *Journal of Chemical Education, 87*, 1438–1443.

Cain, M. S., Leonard, J. A., Gabrieli, J. D., & Finn, A. S. (2016). Media multitasking in adolescence. *Psychonomic Bulletin & Review*, *23*, 1932–1941.

Cameron, P., & Giuntoli, D. (1972). Consciousness sampling in the college classroom or Is anybody listening? *Intellect, 101*, 63–64.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006, September 1). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*, 1307–1310.

Cowan, N., Elliot, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*, 42–100.

Cronbach, L. J., (1957). The two disciplines of scientific psychology. *American Psychologist, 12*, 671–684.

Di Vesta, F. J., & Gray, G. S. (1972). Listening and note taking. *Journal of Educational Psychology, 63*, 8–14.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4–58.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the achiever: The structure of adolscents‘ academic achievement related-beliefs and self-perceptions. *Personality and Social Psychology Bulletin, 21*, 215–225.

Einstein, G. O., Morris, J., & Smith, S. (1985). Note-taking, individual differences, and memory for lecture information. *Journal of Educational Psychology, 77*, 522–532.

Faber, J. E., Morris, J. D., & Lieberman, M. G. (2000). The effect of note taking on ninth grade students' comprehension. *Reading Psychology, 21*, 257–270.

Farley, J., Risko, E. F., & Kingstone, A. (2013). Everyday attention and lecture retention: The effects of time, fidgeting, and mind wandering. *Frontiers in Psychology*, *4*: 619. doi:10.3389/fpsyg.2013.00619

Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review, 20*, 586–592.

Fox, A. B., Rosen, J., & Crawford, M. (2009). Distractions, distractions: does instant messaging affect college students' performance on a concurrent reading comprehension task? *CyberPsychology and Behavior*, *12*, 51–53.

Frank, D. J., Nara, B., Zavagnin, M., Touron, D. R., & Kane, M. J. (2015). Validating older adults' reports of less mind-wandering: An examination of eye movements and dispositional influences. *Psychology and Aging, 30*, 266–278.

Franklin, M. S., Mrazek, M. D., Anderson, C. L., Smallwood, J., Kingstone, A., & Schooler, J. W. (2013). The silver lining of a mind in the clouds: Interesting musings are associated with positive mood while mind-wandering. *Frontiers in Psychology, 4:* 583. doi:10.3389/fpsyg.2013.00583

Fulmer, S. M., D'Mello, S. K., Strain, A., Graesser, A. C. (2015). Interest-based text preference moderates the effect of text difficulty on engagement and learning. *Contemporary Educational Psychology, 41*, 98–110.

Geerlings, T. (1995). Students' thoughts during problem-based small-group discussions. *Instructional Science, 22*, 269–278.

Grodsky, A., & Giambra, L. M. (1990–91). The consistency across vigilance and reading tasks of individual differences in the occurrence of task-unrelated and task-related images and thoughts. *Imagination, Cognition and Personality, 10*, 39–52.

Hadwin, A. F., Kirby, J. R., & Woodhouse, R. A. (1999). Individual differences in notetaking, summarization, and learning from lectures. *Alberta Journal of Educational Research, 45*, 1–17.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical variable. *British Journal of Mathematical and Statistical Psychology, 67*, 451–470.

Hembrooke, H., & Gay, G. (2003). The laptop and the lecture: The effects of multitasking in learning environments. *Journal of Computing in Higher Education, 15*, 46–64.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research, 60,* 549–571.

Hidi, S. (1995). A re-examination of the role of attention in learning from text. *Educational Psychology Review, 7*, 323–350.

Hidi, S., & Reninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111–127.

Hollis, R. B., & Was, C. A. (2016). Mind wandering, control failures, and social media distractions in online learning. *Learning and Instruction, 42*, 104–112.

Hox, J. J. (1998). Multilevel modeling: When and why. In I. Balderjahn, R. Mathar & M. Schader (Eds.), *Classification, data analysis, and data highways* (pp. 147–154). New York: Springer.

Hox, J. J., Moerbeek, M., Kluytmans, A., & van de Schoot, R. (2014). Analyzing indirect effects in cluster randomized trials. The effect of estimation method, number of groups and group sizes on accuracy and power. *Frontiers in Psychology, 5*: 78. doi:10.3389/fpsyg.2014.00078

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology, 102*, 880–895.

Immordino-Yang, M. H., Christodoulou, J. A., & Singh, V. (2012). Rest is not idleness: Implications of the brain's default mode for human development and education. *Perspectives on Psychological Science, 7*, 352–364.

Jing, H. G., Szpunar, K. K., & Schacter, D. L. (2016). Interpolated testing influences focused attention and improves integration of information during a video-recorded lecture. *Journal of Experimental Psychology: Applied, 22*, 305–318.

Junco, R., & Cotton, S. R. (2012). No A 4 U: The relationship between multitasking and academic performance. *Computers & Education, 59*, 505–514.

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

Kane, M. J., & McVay, J.C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science, 21*, 348–354.

Kane, M. J., Meier, M. E., Smeekens, B. A., Gross, G. M., Chun, C. A., Silvia, P. J., & Kwapil, T. R. (2016). Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *Journal of Experimental Psychology: General, 145*, 1017–1048.

Kiewra, K. A., & Benton, S. L. (1988). The relationship between information-processing ability and notetaking. *Contemporary Educational Psychology, 13*, 33–44.

Killingsworth, M. A., & Gilbert, D. T. (2010, November 12). A wandering mind is an unhappy mind. *Science, 330,* 932.

Kintsch, W. (1980). Learning from text, levels of comprehension, or: Why anyone would read a story anyway. *Poetics, 9*, 87–98.

Klinger, E. (1971). *Structure and functions of fantasy*. New York: Wiley-Interscience.

Klinger, E. (2013). Goal commitments and the content of thoughts and dreams: Basic principles. *Frontiers in Psychology, 4:* 415. doi:10.3389/fpsyg.2013.00415

Kobayashi, K. (2005). What limits the encoding effect of note-taking? A meta-analytic examination. *Contemporary Educational Psychology, 30*, 242–262.

Kraushaar, J. M., & Novak, D. C. (2010). Examining the effects of student multitasking with laptops during the lecture. *Journal of Information Systems Education*, *21*, 241.

Lindquist, S. I., & McLean, J. P. (2011). Daydreaming and its correlates in an educational environment. *Learning and Individual Differences, 21*, 158–167.

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement, 70*, 647–671.

Locke, L. F., & Jensen, M. K. (1974). Thought sampling: A study of student attention through self-report. *The Research Quarterly, 45*, 263–275.

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality, 48*, 61–83.

McIntyre, S. (1992). Lecture notetaking, information-processing, and academic achievement. *Journal of College Reading and Learning, 25*, 7–17.

McMillan, R. L., Kaufman, S. B., & Singer, J. L. (2013). Ode to positive constructive daydreaming. *Frontiers in Psychology, 4*: 626. doi:10.3389/fpsyg.2013.00626

McVay, J. C., & Kane, M. J. (2012a). 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.

McVay, J. C., & Kane, M. J. (2012b). 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.

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010, November 26). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science, 330*, 1234–1237.

Moisala, M., Salmela, V., Hietajärvi, L., Salo, E., Carlson, S., Salonen, O., ... & Alho, K. (2016). Media multitasking is associated with distractibility and increased prefrontal activity in adolescents and young adults. *NeuroImage*, *134*, 113–121.

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science, 25*, 1159–1168.

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences, USA, 106*, 15583–15587.

Pekrun, R., Goetz, T., Daniels, L. M., Stupinsky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control-value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology, 102*, 531–549.

Peper, R. J., & Mayer, R. E. (1978). Note taking as a generative activity. *Journal of Educational Psychology, 70*, 514–522.

Peverly, S. T., Ramaswamy, V., Brown, C., Sumowski, J., Alidoost, M., & Garner, J. (2007). What predicts skill in lecture note taking? *Journal of Educational Psychology, 99*, 167–180.

Peverly, S. T., Vekaria, P. C., Reddington, L. A., Sumoski, J. F., Johnson, K. R., & Ramsay, C. M. (2013). The relationship of handwriting speed, working memory, language comprehension and outlines to lecture note-taking and test-taking among college students. *Applied Cognitive Psychology, 27*, 115–126.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*, 667–686.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33–40.

Ravizza, S. M., Uitvlugt, M. G., & Fenn, K. M. (2017). The negative effects of laptop internet use during class. *Neuroscience Letters, 637*, 44–49.

Randall, J. G., Oswald, F. L., & Beier, M. E. (2014). Mind-wandering, cognition, and performance: A theory-driven meta-analysis of attention regulation. *Psychological Bulletin, 140*, 1411–1431.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*, 164–171.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 353–387.

Risko, E. F., Anderson, N., Sarwal, A., Engelhartdt, M., & Kingstone, A. (2011). Everyday attention: Variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology, 26*, 234–242.

Risko, E. F., Buchanan, D., Medimorec, S., & Kingstone, A. (2013). Everyday attention: Mind wandering and computer use during lectures. *Computers & Education, 68*, 275–283.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*, 261–288.

Ruby, F. J. M., Smallwood, J., Sackur, J., & Singer, T. (2013). Is self-generated thought a means of social problem solving? *Frontiers in Psychology, 4*: 962. doi:10.3389/fpsyg.2013.00962

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist, 26*, 299–323.

Schoen, J. R. (1970). Use of consciousness sampling to study teaching methods. *Journal of Education Research, 63*, 387–390.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*, 609–612.

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

Shukor, S. (2005). *Insights into students' thoughts during problem based learning small group discussions and traditional tutorials*. Unpublished manuscript. Retrieved March 18, 2016 from: http://www.tp.edu.sg/staticfiles/TP/files/centres/pbl/pbl_suriya_shukor.pdf

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *Dialogue: The Official Newsletter of the Society for Personality and Social Psychology, 26,* 4–7.

Singer, J. L. (1966). *Daydreaming: An introduction to the experimental study of inner experience*. New York: Random House.

Smallwood, J., & Andrews-Hanna, J. (2013). Not all minds that wander are lost: The importance of a balanced perspective on the mind-wandering state. *Frontiers in Psychology, 4*:441. doi:10.3389/fpsyg.2013.00441

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

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology, 66*, 487–518.

Snow, R. (1989). Aptitude-treatment interaction as a framework for research on individual differences in learning. In P. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and Individual Differences* (pp. 13–59). New York: W. H. Freeman.

Song, X., & Wang, X. (2012). Mind wandering in Chinese daily lives – An experience sampling study. *PLoS ONE, 7*, e44423. doi:10.1371/journal.pone.0044423

Spencer, S. J., Steele, C., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4–28.

Steele, C. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613–629.

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences, 110*, 6313–6317.

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, and Cognition, 39*, 832–842.

Unsworth, N., & McMillan, B. D. (2014). Similarities and differences between mind-wandering and external distraction: A latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta Psychologica, 150*, 14–25.

Varao Sousa, T. L., Carriere, J. S. A., & Smilek, D. (2013). The way we encounter reading material influences how frequently we mind wander. Frontiers in Psychology, 4: 992. doi:10.3389/fpsyg.2013.00892

Wammes, J. D., Boucher, P. O., Seli, P., Cheyne, J. A., & Smilek, D. (2016a). Mind wandering during lectures I: Changes in rates across an entire semester. *Scholarship of Teaching and Learning in Psychology, 2*, 13–32.

Wammes, J. D., Seli, P., Cheyne, J. A., Boucher, P. O., & Smilek, D. (2016b). Mind wandering during lectures II: Relation to academic performance. *Scholarship of Teaching and Learning in Psychology, 2*, 33–48.

Wei, F. Y. F., Wang, Y. K., & Klausner, M. (2012). Rethinking college students' self-regulation and sustained attention: Does text messaging during class influence cognitive learning? *Communication Education*, *61*, 185–204.

Appendix A. *Descriptive statistics for outcome and predictor measures, by research site*

	Research Site								**Site Differences**		
	Site A (*n* = 89)				**Site B** (n = 93)						
Measure	***M***	***SD***	**Min**	**Max**	***M***	***SD***	**Min**	**Max**	***t***	***p***	***d* [95% CI]**
Outcomes											
Mind Wandering Rate	0.36	0.19	0.00	0.85	0.54	0.28	0.00	1.00	**-5.22**	**<.001**	-0.77 [-1.07, -0.47]
Posttest (Z-score aggregate)	0.39	0.77	-1.49	1.61	-0.37	0.76	-1.68	1.45	**6.73**	**<.001**	1.00 [0.69, 1.31]
Situational Interest	2.90	0.70	1.00	4.46	2.61	0.71	1.00	4.50	**2.75**	**.007**	0.41 [0.11, 0.70]
Predictors											
Pretest	0.28	0.14	0.00	0.60	0.23	0.13	0.00	0.50	**2.32**	**.021**	0.35 [0.05, 0.64]
Number of Math Courses	4.39	0.82	3.00	6.00	4.41	0.81	3.00	6.00	-0.13	.899	-0.02 [-0.31, 0.27]
Math Interest	3.04	0.93	1.00	4.78	2.79	0.99	1.11	4.67	1.81	.072	0.27 [-0.03, 0.56]
Learning Confidence	6.24	1.80	2.00	10.00	6.02	2.28	0.00	10.00	0.70	.482	0.10 [-0.19, 0.40]
Incremental Beliefs	3.68	0.91	1.50	5.00	3.46	0.94	1.00	5.00	1.59	.114	0.24 [-0.06, 0.53]
Media Multitasking Habits	2.58	0.95	1.00	4.00	2.73	1.09	1.00	5.00	-0.97	.335	-0.14 [-0.44, 0.15]
Note-Taking Habits	3.69	0.43	2.33	4.56	3.59	0.42	2.44	4.56	1.63	.104	0.24 [-0.05, 0.54]
WMC	0.69	0.18	0.09	1.00	0.63	0.19	0.12	1.00	**2.24**	**.026**	0.33 [0.04, 0.63]

Notes. WMC = working memory capacity. *d* = Cohen's *d*, used as an effect size estimate for each Site A vs. Site B contrast. CI = confidence interval. Degrees of freedom (*df*s) for the t tests = 180. Statistically significant *t* values and their corresponding *p* values are printed in bold.

Appendix B. *Bivariate correlations among all predictor and outcome variables for each research site (Site A correlations below the diagonal and Site B correlations above the diagnonal)*

Measure	1	2	3	4	5	6	7	8	9	10	11
Outcomes											
1. Mind Wandering Rate	--	**-.40**	**-.56**	-.13	-.12	**-.26**	-.05	**-.34**	**.29**	**-.21**	.00
2. Posttest	**-.38**	--	**.26**	**.32**	.08	**.28**	.02	.14	-.13	-.07	**.23**
3. Situational Interest	**-.51**	**.27**	--	-.05	.06	**.47**	.19	**.32**	**-.26**	**.20**	-.06
Predictors											
4. Pretest	**-.26**	**.39**	.20	--	.07	.16	.14	.05	-.02	.05	.15
5. Number of Math Courses	-.17	.03	.01	.07	--	.10	.04	.17	-.02	.10	.10
6. Math Interest	**-.21**	.20	**.27**	.16	**.32**	--	**.32**	**.41**	-.15	.07	.16
7. Learning Confidence	-.12	**.30**	.08	.20	.01	**.36**	--	.03	-.04	.13	.06
8. Incremental Beliefs	-.01	.01	.05	.19	-.01	.10	.11	--	**-.25**	.04	.13
9. Media Multitasking Habits	.09	.03	-.09	.02	.14	-.05	-.02	-.20	--	-.19	.02
10. Note-Taking Habits	.07	.07	-.03	-.01	**.23**	.08	.03	-.02	.04	--	**-.35**
11. WMC	-.18	.02	-.10	.16	.11	-.10	.05	.08	.05	.03	--

Notes. WMC = working memory capacity. Degrees of freedom (*dfs*) = 87 for Site A (*n* = 89) and 91 for Site B (*n* = 93). Significant correlations are printed in bold. Correlations for Site A and Site B are presented below and above the diagonal, respectively.

Appendix C. *Results of regression analyses for the three outcome measures that include research site (Site A vs. Site B) as an additional fixed-effect variable, fully crossed with the experimental (note-taking condition) and individual-differences predictor variables*

Predictor	Mind Wandering				Posttest				Situational Interest			
	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Mind Wandering	---	---	---	---	**-0.26**	0.06	-4.42	**< .001**	**-0.48**	0.08	**-6.36**	**<.001**
Site	**-0.54**	0.14	-3.72	**< .001**	**0.50**	0.11	4.54	**< .001**	0.01	0.14	0.08	.935
Pretest	**-0.15**	0.07	-1.99	**.049**	**0.19**	0.05	3.45	**< .001**	-0.05	0.07	-0.72	.476
Number of Math Courses	-0.11	0.08	-1.51	.135	-0.04	0.06	-0.72	.471	-0.09	0.07	-1.22	.225
Math Interest	-0.15	0.08	-1.87	.063	0.09	0.06	1.54	.125	**0.28**	0.07	3.89	**<.001**
Learning Confidence	0.05	0.08	0.57	.570	0.03	0.06	0.56	.577	0.02	0.07	0.28	.777
Incremental Beliefs	-0.04	0.08	-0.56	.575	-0.10	0.06	-1.89	.062	0.00	0.07	0.00	.999
Media Multitasking Habits	**0.16**	0.07	2.30	**.023**	-0.03	0.05	-0.59	.559	-0.03	0.07	-0.44	.663
Note-Taking Habits	-0.03	0.08	-0.37	.712	-0.00	0.06	-0.08	.934	0.01	0.07	0.10	.921
WMC	-0.05	0.08	-0.73	.469	0.01	0.05	0.14	.890	-0.11	0.07	-1.61	.109
Condition	0.00	0.14	-0.01	.995	0.17	0.10	1.61	.110	0.11	0.13	0.82	.415
Condition x Pretest	**0.31**	0.15	2.10	**.038**	0.19	0.11	1.79	.075	0.01	0.14	0.11	.913
Condition x Number of Math Courses	**0.31**	0.15	2.05	**.042**	-0.10	0.11	-0.89	.378	-0.05	0.14	-0.36	.719
Condition x Math Interest	0.09	0.16	0.55	.587	0.17	0.11	1.45	.150	-0.23	0.14	-1.63	.105
Condition x Learning Confidence	-0.25	0.16	-1.57	.119	0.07	0.12	0.57	.573	0.21	0.15	1.48	.140
Condition x Incremental Beliefs	-0.19	0.15	-1.27	.208	0.19	0.11	1.71	.089	0.09	0.14	0.64	.523
Condition x Media Multitasking Habits	-0.04	0.14	-0.30	.766	**0.21**	0.10	2.05	**.042**	-0.02	0.13	-0.12	.906
Condition x Note-Taking Habits	0.16	0.15	1.06	.289	-0.03	0.11	-0.29	.771	0.07	0.14	0.54	.593
Condition x WMC	-0.11	0.15	-0.74	.458	0.19	0.11	1.79	.076	-0.02	0.14	-0.17	.863
Site x Condition	0.22	0.29	0.77	.442	-0.05	0.21	-0.26	.796	0.15	0.26	0.58	.566
Site x Pretest	-0.01	0.15	-0.08	.936	0.04	0.11	0.38	.708	**0.33**	0.13	2.46	**.015**
Site x Number of Math Courses	0.07	0.15	0.45	.651	-0.12	0.11	-1.05	.298	-0.10	0.14	-0.71	.480
Site x Math Interest	-0.01	0.16	-0.04	.972	-0.13	0.11	-1.15	.254	-0.08	0.14	-0.56	.579
Site x Learning Confidence	-0.10	0.16	-0.61	.541	**0.27**	0.11	2.38	**.019**	-0.12	0.14	-0.86	.394
Site x Incremental Beliefs	**0.31**	0.15	2.00	**.048**	0.09	0.11	0.80	.427	-0.02	0.14	-0.13	.894
Site x Media Multitasking Habits	-0.13	0.14	-0.91	.365	0.11	0.10	1.03	.304	0.06	0.13	0.50	.622
Site x Note-Taking Habits	0.14	0.15	0.90	.368	0.18	0.11	1.65	.102	-0.05	0.14	-0.35	.729
Site x WMC	-0.17	0.15	-1.13	.260	-0.19	0.11	-1.71	.089	-0.04	0.14	-0.31	.754

(continued on next page)

Predictor	Mind Wandering				Posttest				Situational Interest			
	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Site x Condition x Pretest	-0.22	0.29	-0.74	.461	0.14	0.21	0.65	.515	0.10	0.27	0.39	.699
Site x Condition x Number of Math Courses	-0.27	0.31	-0.89	.375	0.01	0.22	0.07	.945	**-0.62**	0.28	-2.25	**.026**
Site x Condition x Math Interest	-0.07	0.32	-0.23	.822	**0.62**	0.23	2.71	**.008**	-0.03	0.29	-0.12	.904
Site x Condition x Learning Confidence	0.03	0.32	0.10	.917	-0.24	0.23	-1.06	.292	-0.15	0.29	-0.53	.595
Site x Condition x Incremental Beliefs	0.34	0.31	1.10	.273	-0.42	0.22	-1.89	.061	-0.03	0.28	-0.09	.926
Site x Condition x Media Multitasking Habits	-0.02	0.29	-0.06	.951	-0.01	0.21	-0.03	.975	0.27	0.26	1.04	.303
Site x Condition x Note-Taking Habits	-0.14	0.30	-0.46	.643	-0.06	0.22	-0.28	.778	0.48	0.28	1.75	.083
Site x Condition x WMC	0.08	0.30	0.26	.799	0.04	0.22	0.18	.859	-0.14	0.27	-0.53	.599

Note: WMC = working memory capacity. Mind wandering model: total $R^2 = .38$, $F(35, 146) = 2.51$, $p < .001$. Posttest model: total $R^2 = .55$, $F(36, 145) = 5.00$, $p < .001$. Situational interest model: total $R^2 = .49$, $F(36, 145) = 3.84$, $p < .001$. The regression models reported in this table are analogous to those reported in Table 4 (for mind wandering) and Table 6 (for posttest scores and situational interest) but include research site (Site) as an additional fixed-effect variable and cross it fully with the note-taking condition (Condition) and seven individual-differences measures, including the three-way interactions involving both Site and Condition. In all of the regression models summarized in this table, the Site variable was coded in the following way: Site B was coded as –0.5 and Site A was coded as + 0.5. Significant coefficients and the corresponding *p* values are printed in bold. The boxes surrounding some entries indicate the predictors that were significant in our primary analyses (i.e., the models without Site and their interactions, summarized in Tables 4 and 6).

Online Supplementary Materials (Tables S1, S2, & S3)

This document includes three tables that summarize the results of the supplementary regression analyses that we conducted separately for each research site (Site A and Site B) for each of the three main outcome variables: mind-wandering (TUT) rate (Table S1), posttest scores (Table S2), and situational interest (Table S3). These site-by-site regression results were intended to supplement the analyses reported in the subsection entitled "Testing the Generalizability of the Main Results Across Two Research Sites" and summarized in Appendix C (see also Appendix A, which presents descriptive statistics for each site, and Appendix B, which presents bivariate correlations among the relevant measures separately for each site).

The models presented in these supplementary tables are analogous to those reported in the main analyses (summarized in Tables 4 and 6), but the analyses were conducted with only the Site A sample (n = 89) or the Site B sample (n = 93), with all the variables entered into the models standardized for each site. Thus, the statistical power for these site-by-site analyses were substantially reduced by splitting the sample roughly in half. To facilitate the comparison between our primary analyses and these site-related supplementary analyses (Appendix C and Tables S1–S3), we indicated those predictors that were significant in our primary analyses (summarized in Tables 4 and 6) by surrounding their entries with boxes.

The results and the take-home messages of these site-by-site analyses are briefly discussed in Footnote 10 of the main article. In summary, although there were some differential effects of some predictors across the two sites (i.e., significant at one site but not at the other), the predictors that were significant in our primary analyses—those indicated by boxes—mostly remained significant (or marginally significant) for both research sites in the posttest (Table S2) and situational interest (Table S3) models. The results were less clear, however, for the mind-wandering model (Table S1), in which none (Site A) or only one (Site B) of the predictors uniquely predicted TUT rate above and beyond the other variables (perhaps in part due to reduced statistical power).

Table S1. *Results of regression analyses for the mind-wandering (TUT rate) outcome variable, conducted separately for each research site*

	Site A (n = 89)				**Site B** (n = 93)			
Predictor	**β**	***SE***	***T***	***p***	**β**	***SE***	***t***	***p***
Pretest	-0.21	0.09	-1.78	.080	-0.12	0.12	-1.17	.246
Number of Math Courses	-0.11	0.09	-0.89	.379	-0.14	0.12	-1.22	.226
Math Interest	-0.19	0.10	-1.53	.130	-0.14	0.12	-1.20	.233
Learning Confidence	0.00	0.11	-0.03	.974	0.10	0.11	0.89	.379
Incremental Beliefs	0.14	0.09	1.21	.231	-0.18	0.12	-1.60	.115
Media Multitasking Habits	0.12	0.09	1.11	.271	**0.23**	0.11	2.13	**.037**
Note-Taking Habits	0.05	0.09	0.45	.655	-0.09	0.12	-0.79	.431
WMC	-0.18	0.09	-1.58	.118	0.03	0.12	0.25	.804
Condition (Note-Taking)	0.18	0.17	0.65	.516	-0.13	0.23	1.72	.090
Condition × Pretest	0.28	0.17	1.17	.248	0.36	0.24	1.74	.086
Condition × Number of Math Courses	0.24	0.18	0.98	.331	0.41	0.24	1.84	.071
Condition × Math Interest	0.06	0.20	0.26	.798	0.11	0.24	0.51	.615
Condition × Learning Confidence	-0.27	0.22	-1.08	.286	-0.27	0.21	-1.25	.214
Condition × Incremental Beliefs	-0.03	0.18	-0.14	.890	-0.34	0.25	-1.48	.144
Condition × Media Multitasking Habits	-0.06	0.18	-0.29	.775	-0.03	0.22	-0.16	.876
Condition × Note-Taking Habits	0.12	0.18	0.51	.614	0.21	0.24	0.95	.345
Condition × WMC	-0.09	0.18	-0.42	.678	-0.14	0.24	-0.62	.540

Note. WMC = working memory capacity. Site A model: total R^2 = .21, $F(17, 71)$ = 1.13, p = .344. Site B model: total R^2 = .31, $F(17, 75)$ = 2.02, p = .020. Significant coefficients and the corresponding p values are printed in bold. The boxes surrounding some entries indicate the predictors that were significant in our primary analyses (i.e., the models without Site and their interactions, summarized in Tables 4 and 6).

Table S2. *Results of regression analyses for the posttest outcome variable, conducted separately for each research site*

	Site A (n = 89)				**Site B** (n = 93)			
Predictor	**β**	***SE***	***T***	***p***	**β**	***SE***	***t***	***p***
Mind Wandering	**-0.36**	0.10	-3.72	**<.001**	**-0.30**	0.08	-2.78	**.007**
Pretest	**0.25**	0.07	2.58	**.012**	**0.21**	0.08	2.21	**.030**
Number of Math Courses	-0.14	0.08	-1.39	.168	0.03	0.08	0.32	.748
Math Interest	0.01	0.08	0.09	.927	**0.22**	0.08	2.06	**.043**
Learning Confidence	0.19	0.09	1.86	.067	-0.16	0.07	-1.59	.116
Incremental Beliefs	-0.06	0.08	-0.62	.538	-0.19	0.08	-1.70	.094
Media Multitasking Habits	0.04	0.08	0.44	.663	-0.14	0.07	-1.35	.182
Note-Taking Habits	0.12	0.08	1.20	.235	-0.12	0.08	-1.13	.263
WMC	-0.13	0.08	-1.34	.186	0.13	0.08	1.25	.215
Condition (Note-Taking)	0.36	0.14	1.07	.288	0.17	0.15	1.32	.193
Condition × Pretest	0.38	0.15	1.96	.054	0.12	0.16	0.63	.529
Condition × Number of Math Courses	-0.10	0.15	-0.47	.637	-0.17	0.16	-0.81	.420
Condition × Math Interest	**0.60**	0.16	2.91	**.005**	-0.20	0.16	-0.96	.341
Condition × Learning Confidence	-0.09	0.18	-0.44	.661	0.29	0.14	1.45	.152
Condition × Incremental Beliefs	-0.03	0.15	-0.15	.885	**0.57**	0.16	2.60	**.011**
Condition × Media Multitasking Habits	0.25	0.15	1.35	.182	0.31	0.14	1.55	.126
Condition × Note-Taking Habits	-0.07	0.15	-0.35	.725	-0.02	0.16	-0.09	.930
Condition × WMC	0.26	0.15	1.39	.169	0.25	0.16	1.16	.251

Note. WMC = working memory capacity. Site A Model: total R^2 = .48, F(18, 70) = 3.59, p < .001. Site B Model: total R^2 = .42, F(18, 74) = 2.94, p < .001. Significant coefficients and the corresponding p values are printed in bold. The boxes surrounding some entries indicate the predictors that were significant in our primary analyses (i.e., the models without Site and their interactions, summarized in Tables 4 and 6).

Table S3. *Results of regression analyses for the situational interest outcome variable, conducted separately for each research site*

Predictor	Site A (n = 89)				Site B (n = 93)			
	β	***SE***	***t***	***p***	**β**	***SE***	***t***	***p***
Mind Wandering	**-0.50**	0.13	-4.88	**<.001**	**-0.42**	0.09	-4.40	**<.001**
Pretest	0.10	0.10	0.92	.360	**-0.19**	0.09	-2.16	**.034**
Number of Math Courses	-0.15	0.10	-1.45	.151	-0.02	0.09	-0.24	.811
Math Interest	0.21	0.11	1.95	.056	**0.35**	0.09	3.60	**.001**
Learning Confidence	-0.04	0.12	-0.35	.730	0.08	0.08	0.91	.367
Incremental Beliefs	0.01	0.10	0.09	.929	0.03	0.10	0.28	.778
Media Multitasking Habits	0.02	0.10	0.20	.843	-0.09	0.09	-0.96	.341
Note-Taking Habits	-0.01	0.10	-0.10	.918	0.04	0.09	0.42	.676
WMC	-0.15	0.10	-1.55	.126	-0.09	0.09	-0.98	.331
Condition (Note-Taking)	0.22	0.19	1.06	.292	0.06	0.18	0.23	.818
Condition × Pretest	0.11	0.19	0.52	.606	-0.07	0.19	-0.40	.693
Condition × Number of Math Courses	-0.35	0.20	-1.63	.107	0.22	0.19	1.16	.251
Condition × Math Interest	-0.24	0.22	-1.11	.272	-0.23	0.18	-1.23	.221
Condition × Learning Confidence	0.09	0.24	0.41	.684	0.35	0.16	1.93	.058
Condition × Incremental Beliefs	0.07	0.20	0.36	.723	0.14	0.19	0.71	.482
Condition × Media Multitasking Habits	0.11	0.20	0.55	.584	-0.16	0.17	-0.89	.377
Condition × Note-Taking Habits	0.34	0.20	1.64	.105	-0.19	0.19	-1.00	.322
Condition × WMC	-0.11	0.20	-0.55	.586	0.06	0.19	0.33	.743

Note. WMC = working memory capacity. Site A Model: total R^2 = .42, $F(18, 70)$ = 2.83, p < .001. Site B Model: total R^2 = .53, $F(18, 74)$ = 4.59, p < .001. Significant coefficients and the corresponding p-values are printed in bold. The boxes surrounding some entries indicate the predictors that were significant in our primary analyses (i.e., the models without Site and their interactions, summarized in Tables 4 and 6).