Metacognitive study strategies in a college course and their relation to exam performance

Cristina D. Zepeda¹ · Timothy J. Nokes-Malach²

Accepted: 30 September 2020 / Published online: 4 November 2020
© The Psychonomic Society, Inc. 2020

Abstract

Several strands of prior work have evaluated students' study strategies and learning activities. In this work, we focus on integrating two of those strands. One has focused on student self-reports of their study practices from a cognitive psychology perspective. The other has focused on classifying student learning activities from a learning sciences perspective using the Interactive, Constructive, Active, and Passive (ICAP) framework (Chi & Wylie, 2014). The current study aims to integrate these two strands of research by testing the implications of the ICAP framework with students' self-reports in a classroom context. Another goal was to address the measurement limitations of the metacognitive study strategy literature by using assessment-specific self-reports with both closed and open-ended questions. Across three noncumulative exams, 342 undergraduates self-reported their study practices before each exam. We then categorized their strategies as either active or constructive in alignment with the ICAP framework. Next, we examined whether these strategies were related to each other and then tested the hypothesis that constructive strategies would be positively associated with better exam performance than active strategies. Students reported using a variety of study practices in which a few active strategies were related to constructive strategies, but constructive strategies were more likely to be related to each other. Lastly, supporting the ICAP framework, many of the constructive strategies were positively related to exam performance, whereas the active strategies were not. This work provides insight into the measurement of students' study strategies and their relations to each other and learning outcomes.

Keywords Learning · Normative practices · Performance · Study strategies · Time management

Students can use a variety of study and time-management strategies while learning or preparing for an assessment. To understand the types of strategies students use, two separate strands of prior work in cognitive psychology and the learning sciences have used different methodologies. One has examined self-reported study activities from a cognitive psychology perspective and has provided a rich understanding of how students perceive their strategy use and the effectiveness of those strategies, revealing students' metacognitive study strategies (Hartwig & Dunlosky, 2012; Karpicke, Butler, &

Roediger, 2009; Kornell & Bjork, 2007; McCabe, 2011; Morehead, Rhodes, & Delozier, 2016; Susser & McCabe, 2013; Wissman, Rawson, & Pyc, 2012). The other has examined students' overt learning activities during specific instructional events from a learning sciences perspective and generated the Interactive, Constructive, Active, and Passive (ICAP) framework to categorize observable learning activities based on their hypothesized cognitive processing and learning outcomes (Chi, 2009; Chi & Wylie, 2014; Menekse, Stump, Krause, & Chi, 2013). The current study aims to build upon and integrate these two strands of research to test whether students' self-reported study strategies, as categorized by the ICAP framework, also produce the predicted learning associations in an authentic classroom context. This integration not only examines an extension of the ICAP framework's assessment and methodology but also expands the metacognitive study strategy literature to include additional strategies that have not been included in prior work.

Both strands of prior research have provided important insights into students' strategy use. For example, despite the large research literature revealing which strategies are more

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13421-020-01106-5>.

✉ Cristina D. Zepeda
cdzepeda@wustl.edu

¹ Department of Education, Washington University in St. Louis, One Brookings Drive, Campus Box 1183, St. Louis, MO 63130, USA

² Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, USA

beneficial for learning and transfer (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Koedinger, Booth, & Klahr, 2013), students typically do not know which strategies are more effective than others and tend to use less effective ones (Hartwig & Dunlosky, 2012; Karpicke et al., 2009; Kornell & Bjork, 2007; McCabe, 2011; Morehead et al., 2016; Susser & McCabe, 2013; Wissman et al., 2012). For instance, students tend to reread text or their notes versus quizzing themselves (Karpicke et al., 2009), or they choose to cram the day before the exam instead of spacing out their study (Susser & McCabe, 2013). These findings highlight the importance of examining students' awareness and beliefs about their normative strategy use. In this work, we examined two aspects of students' metacognitive study strategies, how and when they study.

Processing: How they study

When students study, they can engage in different learning activities, which have implications for the degree to which they are cognitively processing the material. Cognitive engagement, as broadly defined in the motivational literature, occurs when *learners employ* or have the *intent to employ cognitive strategies* to strengthen their learning and performance (Fredricks, Blumenfeld, & Paris, 2004; Wang & Degol, 2014). Chi and Wylie's (2014) definition focuses on operationalizing the different ways a student can engage. They define *cognitive engagement* as "*the way a student engages with the learning materials in the context of an instructional or learning task, reflected in the overt behavior the student exhibits while undertaking an activity*" (pp. 219). To distinguish between different types of cognitive engagement, Chi and colleagues proposed the ICAP framework (Chi, 2009; Chi & Wylie, 2014; Menekse et al., 2013). This framework is useful in that it connects theories of cognition and metacognition to observable learning activities and their hypothesized learning outcomes. To provide support for their framework, the authors reviewed experimental studies that showed results consistent with the framework's predictions. By extending the ICAP framework to self-reports, we evaluate whether the predictions of the framework occur not just when the activity is observable but when people report about their activities. Are these observable activities evident to learners, are they categorizable by the ICAP framework, and do they relate to learning outcomes as predicted by ICAP?

According to the ICAP framework, a learning activity falls into one of four modes or categories of cognitive engagement, with each category having different implications for how well the information is learned and applied. From lowest to highest, these categories include passive, active, constructive, and interactive. Given that students likely engage in multiple learning activities when studying for exams, the strategies they use

likely fall into different categories of engagement. Some might be active while others are constructive. Although the ICAP framework has four categories, due to the constraints of our open-ended data,¹ we only focus on active and constructive.

To apply the ICAP framework to metacognitive study strategies, we surveyed the literature to select a variety of strategies that fit within the active and constructive distinctions (for reviews, see Dunlosky et al., 2013; Karpicke et al., 2009; Koedinger et al., 2013; Richey & Nokes-Malach, 2015). The active category refers to learning activities in which learners attend to and manipulate information (Chi & Wylie, 2014). By attending to and manipulating information, learners are hypothesized to "activate a body of knowledge that is relevant" and "assimilate or integrate new information into the activated schema," making their schemas more complete (Chi & Wylie, 2014, p. 226). For instance, if a student is highlighting information that they think is important or familiar, they are hypothesized to activate relevant prior knowledge structures (e.g., schema) and thereby create an opportunity to add new information provided in the learning resource to their prior knowledge. Chi and colleagues refer to this knowledge acquisition process as gap-filling. Study strategies that support active engagement involve attending to and manipulating information, but not generating new information beyond what the material presents. Active strategies include rewriting (copying notes or material), highlighting (marking the material), and summarizing (paraphrasing and stating the gist of the material).

The constructive category refers to learning activities in which learners generate or produce additional products or ideas that go beyond the information presented in the learning resources (Chi & Wylie, 2014). By generating new information, learners are hypothesized to engage in cognitive processes that infer new knowledge, which includes revising, repairing, reorganizing, and reflecting on one's knowledge (see Chi & Wylie, 2014). Constructive strategies include generating examples (connecting a concept to a situation or case), self-explaining (explaining the material to themselves), analogically comparing (identifying similar features between two concepts, examples, or cases), and quizzing/self-testing (practice retrieving information from memory). We also include two additional constructive strategies that are not usually examined within the metacognitive study strategy literature, but are mentioned in the ICAP literature (Chi & Wylie, 2014) and

¹ Students rarely reported passive strategies that involve receiving information (e.g., listening to lecture) in their open-ended responses likely due to the perception that studying requires some activity outside of class. Students also did not report detailed information about interactive strategies that involve dialogue with another person(s) about the information in which learners are jointly creating and incorporating each other's contributions. For example, students would state they worked with others, but they did not provide details on how they worked with them.

are commonly evaluated within self-regulated learning research (Winne & Perry, 2000; Zimmerman, 2008): metacognitive monitoring (the awareness of one's understanding) and regulation (controlling one's thoughts).

Although monitoring is not usually included as a metacognitive study strategy in and of itself, it is often included as a component in other constructive strategies. For example, one of the conditions under which students are hypothesized to engage in spontaneous self-explanation is when they experience confusion or do not understand some new material, which suggests a form of monitoring (Aleven & Koedinger, 2002; Chi, 2000). Similarly, the self-testing study strategy is often used as a tool to identify whether a student can correctly retrieve the information and provides a concrete way for students to assess and monitor what they know (Karpicke et al., 2009). Theories of comparison and analogical reasoning also appear to require monitoring skills to remember past relevant examples, align features across cases, and generate new inferences based on those alignments (Alfieri, Nokes-Malach, & Schunn, 2013; Gentner, 1983; Gick & Holyoak, 1983). Some of the other constructive strategies have also been hypothesized to be interactive (e.g., self-explanation and comparison; Edwards, Williams, Gentner, & Lombrozo, 2019; Rittle-Johnson & Star, 2011), suggesting that the use of one constructive strategy supports the use of others. Because monitoring plays an important role in many of the constructive study strategies, we coded for it separately as it may play a particularly powerful role in learning.

Several hypotheses can be derived from the ICAP framework about study strategies and student learning. Within the ICAP framework, the categories are hierarchically organized with higher categories inheriting the characteristics of the lower category (e.g., constructive also entails the attending processes that occur within active). This structure suggests that strategies supporting constructive engagement might also be related to strategies that support active engagement, as they both share attending processes. The ICAP framework also posits that active engagement is related to a “shallow” or surface-level understanding in which a learner can apply what they learned to similar situations and problems. Constructive engagement is predicted to be related to a “deeper” or conceptual understanding in which the learner has the potential to transfer that knowledge to novel situations. Prior work has supported these hypotheses such that constructive learning activities resulted in better learning and transfer of information than active learning activities in both classroom and laboratory experiments (Coleman, Brown, & Rivkin, 1997; Menekse et al., 2013). In this prior work, the researchers measured students' engagement by examining evidence of their overt learning activities (e.g., examining whether they self-explained) and examined their relation to

performance. In the current work, we tested whether these findings can be extended to students' self-reported study activity and exam performance. If students' self-reports also show that constructive strategies are more likely than active strategies to be related to learning and performance outcomes, then these results would provide additional evidence for the ICAP framework. It would also contribute a useful and practical set of tools that can be added to the repertoire of measures assessing students' learning practices.

This work also broadens the scope of study strategies within the metacognitive study strategy literature to include additional strategies that have not been previously reported (e.g., Hartwig & Dunlosky, 2012; Karpicke et al., 2009; Morehead et al., 2016). These additional strategies include self-explanations (Chi, De Leeuw, Chiu, & Lavancher, 1994; Nokes-Malach, VanLehn, Belenky, Lichtenstein, & Cox, 2013; Nokes, Hausmann, VanLehn, & Gershman, 2011; Rittle-Johnson, 2006), metacognitive monitoring (Nietfeld, Cao, & Osborne, 2006; Rawson, O'Neil, & Dunlosky, 2011; Tobias & Everson, 2002), and analogical comparison (Alfieri et al., 2013; Gentner, Loewenstein, Thompson, & Forbus, 2009; Nokes-Malach et al., 2013). While the ICAP framework contributes additional strategies to the metacognitive study strategy literature, the metacognitive study strategy literature also includes additional strategies that the ICAP framework does not capture. These additional strategies involve the timing for when students study and are discussed next.

Time management: When they study

Another component of metacognitive study strategy is *when* they choose to study or how they structure their learning. Although students are aware that they should spend time studying, they typically do not view spacing out their study as a “study strategy” (Wissman et al., 2012). Laboratory work has shown that increasing the amount of time students study (Vaughn & Rawson, 2011) and the number of sessions they study (i.e., distributed practice; Rawson & Dunlosky, 2011) results in long-term retention. Other studies have examined how these timing strategies benefit learning when used with retrieval practice (e.g., Kornell, 2009; Pashler, Zarow, & Triplett, 2003; Rawson, Dunlosky, & Sciarrelli, 2013). Although this work has shown that timing can benefit student learning, student reports about their use of cramming (the opposite of spaced practice) and spacing studying sessions were not related to student GPA (Hartwig & Dunlosky, 2012). In this work, we examine student reports of when they study, the amount of time they study, the number of study sessions, and the amount of time they spend on different resources.

Contextual features

In addition to the different types of strategies and methodologies between these two strands of research, they also differ in their contextual features. These differing contextual features include their environmental context and the grain size of strategy use. For example, much of the metacognitive study strategy literature has assessed students' study strategies outside of a particular classroom context (e.g., in a lab environment asking about their general study strategies) and has rarely examined the relation to learning and performance outcomes. Alternatively, the overt study strategy literature has focused on specific learning activities and the relation of those activities to specific learning and performance outcomes in classrooms, educational technology environments, and laboratory studies (Chi & Wylie, 2014). Although there are differences between the approaches of these two strands of research, they both have focused on examining students' strategy use at a single point in time, albeit with different grain sizes of strategy use (general vs. one learning activity). While there are benefits in determining students' general awareness and values of different study strategies and observing their use of strategies during an activity, it is unclear how their strategies occur over time within a particular course context. If students adopt different study strategies for different types of learning objectives and courses, then this variability and nuance would not be captured. To address these limitations, we examined students' metacognitive strategies for a particular set of exams within a specific class context. Asking about study strategies as a part of the course provides a specific context (course content and structure) and time to frame the interpretation and generation of the study strategies students report.

By examining student strategies in a particular educational context, we can also examine whether these self-reports are related to each other and specific performance outcomes in classroom tasks, providing a test of the ICAP framework. Past work in the metacognitive study strategy literature has not often examined how these strategies are related to each other, and only one study has examined the relation between the study strategies and classroom performance (e.g., GPA; Hartwig & Dunlosky, 2012). While this approach reveals important relations to general academic performance, it does not address class-specific relations. GPA is a very coarse-grain measure of performance that consists of multiple class grades, so it is unclear whether certain study strategies are related to specific performance outcomes. For instance, a grade for one course might have different weights (or no weight at all) for exams, assignments, attendance, participation, and extra credit in comparison to a grade for another course, making it difficult to discern the relationship between the strategies and knowledge assessments.

To better align the strategies with knowledge assessments, we examined students' self-reports on the types of strategies

they employed for particular exams across a semester. This approach reveals which strategies students think are critical for studying for an exam and whether there are certain strategies that tend to be used together, which can provide a baseline for future interventions. Further, it allows us to capture the frequency in which students use different types of strategies. Prior work has typically examined either the study activities during one learning event or general study strategies making it difficult to discern students' frequency of use of particular strategies and the relation of this strategy use to other strategies and performance.

Variation in the types of questions

Within the metacognitive study strategy literature, there has also been variation in the types of questions that have been used to measure students' strategy use. Much of the work has asked students to endorse whether they used certain study strategies with closed-ended questions (i.e., prompting specific strategies), instead of asking them to describe how they study in their own words. For example, two studies asked students, "Which of the following study strategies do you use regularly?" and then supplied students with a list of specific strategies (e.g., recopy notes, test yourself with questions or problems; Hartwig & Dunlosky, 2012; Morehead et al., 2016). Other work, outside of the metacognitive study strategy literature, in educational psychology has used Likert-scale measures that prompt students to respond to specific cognitive, metacognitive, and time-management strategies from the Motivated Strategies for Learning Questionnaire (MSLQ; see Credé & Phillips, 2011, for a review). However, these scales have been critiqued for their validity and often include a narrow view of the strategies (Dunn, Lo, Mulvenon, & Sutcliffe, 2012). Although these prompt-based questions provide information for how students view specific strategies, it prevents students from freely describing which strategies they use in their own words, and it potentially biases their responses by providing specific formulations of study strategies that they may not generate on their own. Students may think that a strategy is useful and report using it even though they did not actually use it.

To complement this approach of asking about specific strategies, it might benefit researchers to ask students about their study strategies via an open-ended question. Open-ended questions can reveal additional information about which strategies are most salient to students and how these strategies emerge from students' regular practices. The open-ended question might also reduce the degree to which students overestimate which strategies they use (Kruger & Dunning, 1999), as it allows for the reporting of a study strategy without specific prompting. It can also provide information to differentiate those who view a learning activity as a study strategy from

those who, only when prompted, think they are using it or that it would be good to use. The open-ended approach provides additional affordances for measuring these strategies as it can provide important information as to how a student views their own study strategies, and potentially be a more valid way of assessing their strategies (see Ericsson & Simon, 1980). When students respond to an open-ended prompt, they describe their strategies in their own words, and the experimenter then has to classify the strategy, as opposed to students responding to a study strategy prompt, which requires them to interpret what that strategy means for a given context.

Some prior work provides initial support that these approaches reveal differential information. For instance, when comparing the responses of an open-ended question asked by Karpicke et al. (2009; “What kind of strategies do you use when you are studying?”) to the prompting questions that provided specific strategies asked by Hartwig and Dunlosky (2012) and Morehead et al. (2016), a lower percentage of students reported using some strategies in the open-ended versus the prompting question. For example, 10.7% of the participants in Karpicke et al. (2009) reported using self-testing as compared to 71% of the participants in Hartwig and Dunlosky (2012) and 72% in Morehead et al. (2016). A similar pattern was found for highlighting: 6.2% (Karpicke et al., 2009) versus 72% (Hartwig & Dunlosky, 2012) and 53% (Morehead et al., 2016), respectively. Given that these questions assessed students’ general strategy use outside the context of a specific class, it might have been more difficult for students to retrieve how they have generally studied (cf. reflecting on their use in a specific class), resulting in fewer students reporting those strategies in their open-ended responses. Therefore, to evaluate whether this misalignment occurs when the questions are framed to a specific class, we include both open-ended and strategy-specific prompts to evaluate whether they have converging validity. If these measures converge and have similar predictive validity, then it provides further support for the use of the easier diagnostic tool, the prompting questions.

The current study

The main goal of this study was to integrate prior work on metacognitive study strategies with the ICAP framework to (1) test whether ICAP predictions with students’ self-reported study strategies and their relation to class performance, (2) to expand the types of metacognitive study strategies that have been previously examined, (3) to investigate the convergent validity of two types of self-reports, and (4) to better understand students’ normative study practices in a large lecture course. To do so, we examined students’ reported study practices for three noncumulative exams that assessed factual, conceptual, and applied knowledge. Investigating these strategies over time enabled us to evaluate whether students’

frequency in reporting specific strategies was related to performance in the class as well as whether the strategies were related to each other. To assess these strategies, we used a questionnaire that included an open-ended question in which students were asked to describe their study practices as well as questions that prompted students to reflect on their use of specific strategies. We chose to assess some of these strategies via Likert-type scales or forced-choice items, as they were not captured in prior work using an open-ended question (Karpicke et al., 2009), suggesting that these might be aspects of studying that students are unaware of using. These strategies include regulating and monitoring understanding, using self-explanations, making analogical comparisons, and the spacing of study. With the longitudinal data, we also aimed to evaluate students’ normative practices by examining their frequency of strategy use, the relation between the frequency of use among strategies, and the relation of this use to average exam performance. Specifically, we hypothesized the following:

H1. Open-ended and Likert-scale measures would be positively related to each other and show convergent results.

H2. The more frequently students reported spacing their studying, and using constructive and active strategies across exams, the more likely they would perform better on the exams. Further, more constructive strategies would be positively related to exam performance relative to active strategies.

Method

Participants

Approximately 395 undergraduates participated in the study as part of their normal completion of a psychology course at a large mid-Atlantic university. At this university, this course can serve as part of a general education requirement; therefore, both majors and nonmajors take the course. Students received one extra credit point on their exams if they completed the survey before the exam. Fifty-three students did not complete all the materials ($N = 41$) or were missing demographic data ($N = 12$), resulting in 342 students. Of those 342, 170 undergraduates were enrolled in one semester of the course, and 172 were enrolled in another semester; both were taught by the same instructor, but had different class assignments (warm-up quizzes versus end-of-lecture quizzes, different types of homework). See Table 1 for the sample characteristics.

This sample is of adequate sample size, as prior work revealed effects relating some of these strategies (e.g., self-testing) to GPA with a sample of 324 students ($\gamma = .28, p =$

Table 1. Sample characteristics

Characteristic	<i>M (SD)</i> or percentage
High school GPA	3.91 (0.57)
Age	20.53 (1.42)
Psychology major	70.76%
Female	73.10%
White	71.35%
Asian	17.84%
Black	2.92%
Hispanic/Latinx	0.29%
White and Asian	1.75%
Multiple races/ethnicities with at least 1 traditionally underrepresented race/ethnicity	5.85%

.001; Hartwig & Dunlosky, 2012). We also conducted a power analysis using the WebPower in R in which we entered an alpha of 0.05, power of 0.80, and a small effect size ($r = .20$) for a two-tailed semipartial Pearson's correlation test (we partialled out the variance for six student-level variables which are described below), resulting in the determined sample size to be 199 (Zhang & Yuan, 2018). This sample size is also similar to the suggestions recommended by Bonett and Wright (2000) for a Kendall tau rank correlation without any covariates.

Measures

Questionnaire Students answered open-ended and closed-ended (Likert-scale, forced-choice) questions about their study practices. See Table 2 for all questionnaire items and their scales or selection option if applicable.

The Likert-scale items were adapted from prior work (O'Neil & Abeli, 1996; Pintrich, Smith, Garcia, & McKeachie, 1991; Schraw & Dennison, 1994; Zepeda, 2016) and measured four constructive strategies: monitoring (two items), metacognitive regulation (four items), self-explanation (two items), and comparison (three items). Using the GPArotation (Bernaards & Jennrich, 2005) and *lavaan* (Rosseel, 2012) packages in R, EFA and CFA analyses revealed four separate factors, one for each constructive strategy (EFA for Exam 1: RMSEA = .039, TLI = .976; CFA for Exam 2: RMSEA = .046, SRMR = .039, CFI = .981, TLI = .973; CFA for Exam 3: RMSEA = .051, SRMR = .036, CFI = .978, TLI = .968). Each of these factors had adequate internal reliability across the three time points as reflected by Cohen's alpha or Spearman–Brown's coefficient (appropriate for scales with only two items; see Table 2). We captured these strategies with Likert-scale items because prior work investigating student reported strategies had not captured or reported

student use of these strategies, suggesting that perhaps students did not report using them. These Likert-scale measures would be able to capture whether students only report using these strategies if they are prompted.

While the Likert-scale items were adapted from prior work, other items were developed for this study and contained elements that have been examined in prior work (see Table 2). For example, prior work had examined the open-ended question with a different variation (“What kind of strategies do you use when you are studying?”; Karpicke et al., 2009) and coded this question for four of the 12 strategies we coded (rewriting, highlighting, creating examples, and quizzing). Rewriting, highlighting, cramming, quizzing, and spacing were also captured by Hartwig and Dunlosky (2012) and Morehead et al. (2016), but with a prompting question (“Which of the following study strategies do you use regularly?”).

Open-ended coding. For the open-ended question that asked students to describe how they studied for each exam, we developed and applied a coding protocol ($N = 1,026$ responses, 342 for each exam). See Table 3 for the protocol. The open-ended protocol was developed by reading through a subset of students' reports and evaluating whether there was evidence for a particular study or time-management strategy based on theory. Some strategies (e.g., rereading and metacognitive regulation) were not readily evident or clearly definable given student descriptions, so they were not included in the rubric. Then operational definitions were developed for each of the remaining strategies. These codes captured how students studied the information, which we generally refer to as study strategies, included rewriting, highlighting, summarizing, quizzing/self-testing, generating examples, self-explaining, analogically comparing, and monitoring. (See the [Supplemental Materials](#) for more information about each of these strategies.) Another set of codes captured how students spaced their studying, which we referred to as time-management strategies. Two coders separately coded the data. To code the data, they were instructed to read the statement and to code the study strategies in one pass, and then on a later iteration code the time-management strategies to help the coder stay in the frame of the specific type of study approach. For each code, the strategies were coded as absent (0) or present (1). Coders were also instructed to review and confirm their codes once they finished coding. The two coders had adequate reliability ($\kappa > .7$). See the [Supplemental Materials](#) for example statements for each of the codes.

Exam scores The exams were not cumulative and each consisted of 35 multiple-choice questions that evaluated factual, conceptual, and applied knowledge. Each exam covered

Table 2. Questionnaire items

Question type	Study strategy	Specific strategy	Item
Open-ended Likert scale (1 = <i>strongly disagree</i> to 6 = <i>strongly agree</i>)	Multiple Cognitive - constructive	Self-explanation* ($\rho_{xx'}$ s: .84–.87) Analogical comparison* (α s: .82–.84) Metacognitive monitoring* ($\rho_{xx'}$ s: .65–.73) Metacognitive regulation* (α s: .71–.80)	In a paragraph or two, please describe your study techniques. When I study, I pause to explain to myself difficult concepts so that I am sure I understand them. If I don't understand something, I stop to try to explain it to myself. When I study, I compare and contrast different examples to one another. To better understand one concept, I compare it to another one. When studying class material, I compare and contrast different ideas to one another. I knew when I understood a concept well. It was difficult for me to determine when I knew a concept well. (Rev) The first thing I did was think through a plan for what, when, and how to study. I tried to anticipate difficulties I might have had when learning certain ideas or concepts. I kept track of how well I understood each concept. I kept track of my progress, and if necessary, I changed my study strategies. How much total time did you spend studying in minutes for this EXAM (try to be as specific and as accurate as you can)?
Open-ended	Time management	Amount of time	What percentage of that time was spent reviewing the book? What percentage of that time was spent reviewing your lecture notes/slides? What percentage of that time was spent reviewing the study guide? How many study sessions did you have (regardless of the length of time)?
Force-choice (0, 25, 50, 75, 100)			How often did you read the assigned chapter BEFORE coming to class? How often did you read the assigned chapter AFTER coming to class?
Force-choice (1, 2, 3, 4, 5, 6 or more)		Spacing of study sessions	
Force-choice (1 = <i>never</i> to 5 = <i>always</i>)		Timing of reading	

Note. * indicates that these items were adapted from preexisting measures. All other items were new or developed for this questionnaire. The questionnaire also included additional variables that were beyond the scope of this work, such as motivational measures and questions about how they changed their strategies

Table 3. The protocol

Strategy type	Strategy	Coded if students mentioned
Active	Rewriting	Rewriting their notes or study guide regardless of how much they rewrote
	Highlighting	Highlighting anything related to the course, including the textbook, notes, study guide
	Summarizing	Summarizing anything related to the course, including the textbook, notes, study guide
Constructive	Quizzing/Self-testing	Quizzing themselves using the resources we provide or their own resources
	Examples	Creating their own examples
	Self-explaining	Explaining concepts/topics to themselves
	Analogical comparing	Comparing or contrasting concepts
Time management	Monitoring	Keeping track of their awareness of their understanding, checking to see when they did or did not know something
	Throughout	Studying weeks leading up to the exam/throughout the semester
	A few days	Studying the week or a few days before the exam
	Crammed	Cramming the day before the exam
	Didn't study	That they didn't study

approximately one third of the class material. Before the first exam, students were given example test items so that they would be familiar with the format. The first exam covered history, approaches, methods, perception, and attention. The second exam covered memory (working memory, short-term memory, long-term memory), and visual imagery. The third exam covered language, concepts, problem-solving, expertise, and creativity. Note that for the second exam students learned about retrieval practice and elaboration, but we did not see any increases in the use of these strategies from Exam 1 to Exam 2. Similarly, for the third exam they learned about self-explanation and analogical comparison, and again we did not see any increases in the use of these strategies from Exam 2 to Exam 3.

Procedure

The instructor of the course distributed all of the survey materials via email 3 days before each exam. Students were instructed to complete the survey once they were finished studying for the exam. Within the questionnaire, students first responded to the force-choice and open-ended questions about the amount of time they spent studying. Then they answered the open-ended question about their study practices, and then they responded to Likert-scale questions about the constructive strategies. To receive extra credit, students had to respond to the surveys before the exam. This procedure was repeated for each exam.

Results

To address each of our aims and hypotheses, the results are divided into four sets of analyses. The first set of analyses

examined which strategies students used and how often they used them across the exams to have a better understanding of students' normative study practices. The second set of analyses explored the relations of strategy use within and across the ICAP categories. The third set of analyses evaluated the relation between the open-ended and Likert-scale items that measured the same strategies to test the hypothesis that these measures would be positively related. Lastly, the fourth set of analyses evaluated the relation between the frequency of strategy use and exam performance to test the ICAP hypothesis that more constructive strategies would be positively related to performance than active strategies and to evaluate whether spacing of study practice was positively related to performance. For the correlational analyses, we used Kendall's tau partial correlations with a set of covariates (described below). We also set the alpha level at .05, and report relations for p values less than .05 and marginal relations for p values less than .10 (Keppel & Wickens, 2004). When appropriate, we made Bonferroni corrections to reduce the likelihood of a Type I error. We also converted the Kendall's tau correlations to Cohen's d to be able to interpret the effect sizes of the relations as small when $d < 0.2$, medium when $0.2 < d < 0.8$, and large when $d > 0.8$ (see Cohen, 1988; Olejnik & Algina, 2000; Walker, 2003).

Use and frequency of study and time-management strategies

To have a better understanding of students' normative study practices for each exam and across all the exams, we report a variety of descriptive statistics. For the open-ended responses, we report the percentage of students that described using a specific strategy for each exam, the frequency that a student reported

Table 4. Descriptive statistics for the open-ended study strategies for each exam and across exams

Type	Strategy	Exam 1		Exam 2		Exam 3		Across the exams				Averaged #
								# of times the strategy was mentioned, %				
		<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	0	1	2	3	
Active	Rewriting	87	25	36	11	40	12	66	23	8	3	0.48 (0.76)
	Highlighting	47	14	46	13	38	11	73	18	7	2	0.38 (0.72)
	Summarizing	19	6	11	3	12	4	91	7	2	0	0.12 (0.42)
Constructive	Quizzing/Self-testing	96	28	55	16	49	14	59	27	8	5	0.58 (0.83)
	Creating examples	16	5	15	4	15	4	89	8	2	0	0.13 (0.42)
	Self-explaining	29	8	49	14	43	13	80	17	3	1	0.35 (0.64)
	Analogically comparing	24	7	38	11	20	6	73	20	6	1	0.24 (0.52)
	Monitoring	212	62	189	55	177	52	18	24	28	30	1.69 (1.08)
Avg. # of active	<i>M (SD)</i>	0.45 (0.65)		0.27 (0.50)		0.26 (0.54)						0.33 (0.34)
Avg. # of constructive	<i>M (SD)</i>	1.10 (0.92)		1.01 (0.76)		0.89 (0.83)						1.00 (0.70)

Note. *N* = number of students; % = percentage of students; # = number; Avg. = average. For the rows with the specific strategies; the last column contains the average number of times a strategy was reported across the exams. For the last two rows, the last column contains the average number of strategies for the respective strategy type across the exams

using a strategy across the exams (e.g., the number of times students reported monitoring), and the average number of active and constructive strategies reported for each exam and across the exams. For the Likert-scale and forced-choice responses, these statistics include the average response for each exam and the averaged student response across the exams.

Study strategies: Open-ended question The study strategies included rewriting, highlighting, summarizing, quizzing/self-testing, generating examples, self-explaining, analogically comparing, and monitoring. (See Table 4 for the descriptive statistics for these study strategies.) Students reported using an average of 1.33 strategies for a given exam. They reported that they monitored their understanding the most, followed by quizzing, and rewriting information. Students rarely reported summarizing and creating examples. Monitoring had the largest variation ($SD = 1.08$), whereas students' reports of summarizing and creating examples had the least variation ($SDs = 0.42$). Monitoring was the only study strategy that the majority of students reported using for at least one exam (82%).

When comparing constructive versus active strategy use across the exams, students reported using an average of 0.33 active strategies and 1.00 constructive strategy per exam. We divided this average by the total number of possible strategies for each category (3 for active, 5 for constructive) to calculate a percentage of strategies to compare across ICAP categories. This revealed that students on average reported using 11% of the active strategies and 20% of the constructive strategies for a given exam.

Study strategies: Likert-scale questions The Likert-scale items were on a 6-point rating scale. Results showed that students thought they used a moderate amount of self-explaining, analogically comparing, monitoring, and regulating (see Table 5). These responses were fairly consistent across each exam and had a similar amount of variation.

Time-management strategies: Open-ended question Time-management strategies for the open-ended question involved how students structured the timing of their study, which

Table 5. Descriptive statistics for the study strategy Likert-scale items for each exam and across exams

Strategy	Exam 1		Exam 2		Exam 3		Averaged across exams	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Self-explaining	4.75	0.84	4.76	0.83	4.87	0.80	4.80	0.70
Analogical comparing	4.28	0.88	4.26	0.82	4.36	0.79	4.30	0.67
Monitoring	4.41	0.89	4.39	0.78	4.42	0.79	4.41	0.68
Regulating	4.13	0.78	4.17	0.83	4.33	0.79	4.21	0.70

Table 6. Descriptive statistics for the open-ended time-management strategies for each exam and across exams

Strategy	Exam 1		Exam 2		Exam 3		Across the exams				Averaged #
							# of times the strategy was mentioned, %				
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	0	1	2	3	
Throughout	31	9	22	6	14	4	85	11	4	1	0.20 (0.52)
A few days before	32	9	20	6	23	7	83	13	3	1	0.23 (0.55)
Crammed	2	1	3	1	2	1	98	2	0	0	0.02 (0.14)
Didn't study	0	0	0	0	0	0	100	0	0	0	0 (0)
Avg. # of strategies	<i>M (SD)</i>	0.19 (0.39)	0.13 (0.35)		0.13 (0.33)						0.15 (0.26)

Note. *N* = number of students, % = percentage of students, # = number, and Avg. = average. For the rows with the specific time-management strategies, the last column contains the average number of times a strategy was reported across the exams. For the last row, the last column contains the average number of time-management strategies across the exams

included not studying, cramming the night before, studying a few days before, or studying throughout the semester. (See Table 6 for the descriptive statistics for the open-ended time-management strategies.) Students rarely mentioned timing in their open-ended responses, and none of the students said they did not study. When examining students' frequency in mentioning that they spaced their studying (regardless of length), only 30% of students referenced a timing for at least one exam. Studying throughout and a few days before the exam were the most common reports. In general, time-management strategies were not consistently mentioned across the exams. These results reveal that students do not typically view timing as a study strategy.²

Time-management strategies: Open-ended and forced-choice questions To evaluate time-management strategies, students responded to a subset of questions that asked them about when they read the book, the number of minutes they studied, the percentage of time they reviewed specific materials, and the number of study sessions. (See Table 7 for the descriptive statistics for these measures.) Students sometimes read the book before the class and sometimes after class. When students were asked to enter the amount of time they studied (in minutes), some students responded with data that were not translatable into minutes (e.g., “maybe four days”; “few hours a day”) and were therefore not included in the time analyses. Across the exams, students averaged studying for 363.91 minutes (6 hours and 4 minutes; *SD* = 244.11). Although there was minimal variation in the standard deviations for the amount of time students reported studying across the exams, the standard deviations were large. Students reported spending the majority of their time reviewing the lecture notes/slides.

² Although it might appear that students do not report many strategies, the majority of their responses involved describing the resources they used. Averaged across the exams, the most common resources were notes (92%) and the study guide (75%).

Students also did not space out their studying across the semester as they reported an average of three and half study sessions per exam. The number of study sessions also had minimal variation across the exams.

Relation between and across active and constructive study strategies

Identifying which strategies students think that they used for each exam provides insight into whether these strategies are used in a college course. However, it does not tell us whether the strategies were associated with each other. We examined whether constructive and active strategy use were related to one another as well as whether there were more relations within a category type. For example, are constructive strategies more likely to be positively related to other constructive strategies than active strategies? In addition to providing further insight into the relations within the ICAP framework, this set of analyses could be used to identify a set of interrelated study strategies in an instructional intervention. To evaluate these relations, we used Kendall's tau partial correlations with the *ppcor* package in R to remove the variance explained by student sex, race/ethnicity, age, high school GPA, major, and class section and to reflect the nonnormality of the data (Kim, 2015). This approach removed the variance explained by these covariates and revealed relations above and beyond those variables. Sex (male = 0, female = 1), racial and ethnic representation (represented majority—White and/or Asian = 0; underrepresented minority—Black, Hispanic/Latinx, or Multiple Races/Ethnicities with at least one being underrepresented = 1), major (nonpsychology major = 0, psychology major = 1) and class (fall course = 0, spring course = 1) were dichotomously coded. High school GPA and age were included as continuous variables. Students' open-ended responses were also coded dichotomously (absent = 0, present = 1) and their responses to the Likert-scale measures were coded as

Table 7. Descriptive statistics for time-management strategies for each exam and across exams

		Read the book		Time studying	Percentage of that time reviewing			# of study sessions
		Before class	After class		Book	Lecture notes/slides	Study guide	
Exam		Never (1) to Always (5)		Minutes	0%, 25%, 50%, 75%, 100%			1, 2, 3, 4, 5, 6 or more
Exam 1	<i>M</i>	2.71	2.85	363.61 (<i>N</i> = 340)	26.68%	56.21%	45.69%	3.56
	<i>SD</i>	1.21	1.22	296.92	17.73%	21.49%	27.04%	1.41
Exam 2	<i>M</i>	2.35	2.90	334.46 (<i>N</i> = 339)	26.46%	55.41%	44.17%	3.54
	<i>SD</i>	1.6	1.23	220.67	18.10%	22.00%	27.12%	1.42
Exam 3	<i>M</i>	2.16	2.67	404.09 (<i>N</i> = 339)	21.20%	60.09%	43.42%	3.70
	<i>SD</i>	1.14	1.21	394	17.34%	21.76%	28.22%	1.51
Averaged across exams	<i>M</i>	2.41	2.81	363.91 (<i>N</i> = 336)	24.78%	57.24%	44.42%	3.61
	<i>SD</i>	1.04	1.03	244.11	14.09%	17.46%	23.46%	1.24

Note. *N* = 342 unless otherwise stated

continuous. We also applied Bonferroni corrections to each question type (open-ended and Likert scale). Given that there were 28 correlations among the open-ended strategies, significant relations were reported when $p < .05/28 = .0018$ and marginal relations were reported when $p < .10/28 = .0036$. For the Likert-scale questions there were six correlations such that significant relations were reported when $p < .05/6 = .008$ and marginal relations were reported when $p < .10/6 = .017$.

The correlations between each open-ended study strategy are presented in Table 8. Between active and constructive strategies, there were a few positive relations. Highlighting was positively related to monitoring. Summarizing was positively related to creating examples and comparisons. Within the active strategies, rewriting and highlighting were positively related, but there were no associations to summarizing. For constructive strategies, there were several medium to large positive associations. Self-explaining was positively related to quizzing,

creating examples, comparison, and monitoring. Comparison was positively related to creating examples and quizzing. Monitoring was also positively related to quizzing. For the Likert-scale items, they were all positively related (see Table 9).

Relation between question types: Open-ended and Likert-scale items

Three of the study strategies were measured by both open-ended and Likert-scale items: monitoring, self-explanation, and analogical comparison. To evaluate whether these two types of measures were aligned across the three strategies, we used Kendall's tau partial correlations with the *ppcor* package in R with the same covariates as the earlier analyses: student sex, race/ethnicity, age, high school GPA, major, and class section and to reflect the nonnormality of the data (Kim, 2015). We also applied a Bonferroni correction. Given that

Table 8. Kendall's tau partial correlations between the open-ended study strategies

Strategy Type	Strategy	1	2	3	4	5	6	7	8
Active	1 Rewriting	–	0.16 (0.51)	0.08 (0.25)	0.07 (0.22)	–0.02 (0.06)	–0.01 (0.03)	–0.01 (0.03)	0.09 (0.28)
	2 Highlighting	<.001	–	–0.02 (0.06)	0.09 (0.28)	0.04 (0.13)	0.08 (0.25)	0.08 (0.25)	0.17 (0.55)
	3 Summarizing	.031	.670	–	0.001 (0.003)	0.12 (0.38)	0.08 (0.25)	0.20 (0.65)	0.03 (0.09)
Constructive	4 Quizzing/Self-testing	.057	.012	.975	–	0.09 (0.28)	0.21 (0.68)	0.13 (0.41)	0.24 (0.79)
	5 Creating examples	.554	.280	<.001	.011	–	0.30 (1.02)	0.27 (0.90)	0.03 (0.09)
	6 Self-explaining	.819	.038	.021	<.001	<.001	–	0.26 (0.87)	0.18 (0.58)
	7 Analogically comparing	.728	.023	<.001	<.001	<.001	<.001	–	0.08 (0.25)
	8 Metacognitive monitoring	.018	<.001	.350	<.001	.396	<.001	.024	–

Note. *p* values are provided without the correction in the lower-left diagonal and r_t are provided in the upper left of the diagonal with Cohen's *d* in parentheses. Only those that are $p < .0018$ are significant and are in boldface

Table 9. Kendall's tau partial correlations between the Likert-scale study strategies

Strategy		1	2	3	4
1	Self-explaining	–	0.44 (1.65)	0.33 (1.14)	0.43 (1.60)
2	Analogically comparing	<.001	–	0.21 (0.68)	0.37 (1.31)
3	Metacognitive monitoring	<.001	<.001	–	0.21 (0.68)
4	Metacognitive regulation	<.001	<.001	<.001	–

Note. *p* values are provided without the correction in the lower-left diagonal and r_t are provided in the upper left of the diagonal with Cohen's *d* in parentheses. Only those that are $p < .008$ are significant and are boldface

there were three correlations per exam, significant relations were reported when $p < .05/3 = .017$ and marginal relations were reported when $p < .10/3 = .033$.

The correlations between the statements and their respective Likert-scale measures are presented in Table 10. For Exams 1 and 3, each strategy statement had a medium positive relation to their respective Likert-scale measure. Similar results were obtained for Exam 2, with the exception of monitoring. There was no relation between the monitoring statements and the monitoring Likert-scale measure for Exam 2.

Relation between strategy use and exam performance

In addition to examining whether the study strategies were related to each other, we also examined which strategies were associated with better performance to test the ICAP hypothesis that constructive strategies are more likely than active strategies to be related to performance. To answer this question, we examined the correlations between the number of times students reported a specific strategy (their frequency in reporting a strategy) and their average exam performance. Students generally performed well on the exams with the average exam score in the B range ($M = 82.92$, $SD = 8.08$). For these analyses, we used Kendall's tau partial correlations with the *ppcor* package in R (Kim, 2015) with the same covariates as the earlier analyses: student sex, racial and ethnic representation, age, high school GPA, major, and class section. We also applied a Bonferroni correction. Given that there were 25

correlations, significant relations were reported when $p < .05/25 = .002$ and marginal relations were reported when $p < .10/25 = .004$. All these correlations are presented in Table 11.

Study strategies For the open-ended responses, there were medium positive relations between students' exam performance and stating that they quizzed themselves and monitored their understanding. There were no other relations to exam performance. From these relations, it is evident that more of the constructive strategies were positively related to exam performance than the active strategies, even when accounting for the disproportional number of active ($0/3 = 0\%$ of active strategies were related to exam performance) versus constructive strategies ($2/5 = 40\%$ of strategies were related to exam performance). Similarly, the average number of constructive strategies had a medium positive relation with exam performance, whereas the average number of active strategies did not. Likewise, three of the four Likert-scale constructive strategies had medium positive relations with exam performance.

Time-management strategies For all of the time-management measures, there was no relation to average exam performance. One reason for the lack of relations might be due to the floor effects with the coded open-ended measure of time and the questions about when students read the book. The other measures of time, which includes the number of minutes they studied, the amount of time dedicated to three different resources, and the number of study sessions, did not have floor effects.

Table 10. Kendall's tau partial correlations between the Likert-scale measures and their respective open-ended measure

Likert-scale measure	Exam 1 statement			Exam 2 statement			Exam 3 statement		
	r_t	<i>p</i>	<i>d</i>	r_t	<i>p</i>	<i>d</i>	r_t	<i>p</i>	<i>d</i>
Self-explaining	.19	<.001	0.62	.18	<.001	0.58	.12	.001	0.38
Analogical comparing	.09	.016	0.28	.13	<.001	0.41	.14	<.001	0.45
Monitoring	.10	.007	0.32	.05	.134	0.16	.10	.006	0.32

Note. *p* values are provided without the correction; therefore, only those that are $p < .017$ are significant and are in boldface

Table 11. Relations between each strategy and exam performance by measurement type

Strategy		N	Open-ended measure			Likert-scale measure			Forced-choice measure		
type	Strategy		r_t	p	d	r_t	p	d	r_t	p	d
Active	Rewriting	342	.02	.633	0.05						
	Highlighting	342	.04	.288	0.12						
	Summarizing	342	.05	.160	0.16						
	Number of active strategies	342	.05	.203	0.16						
Constructive	Quizzing/Self-testing	342	.11	.0017	0.36						
	Creating examples	342	.03	.417	0.09						
	Self-explaining	342	.04	.282	0.12	.18	<.001	0.59			
	Analogically comparing	342	.08	.037	0.24	.13	<.001	0.41			
	Metacognitive monitoring	342	.13	<.001	0.40	.14	<.001	0.43			
	Metacognitive regulation	342				.10	.004	0.33			
	Number of constructive strategies	342	.14	<.001	0.46						
Time management	Throughout	342	.03	.403	0.10						
	A few days before	342	.07	.050	0.23						
	Crammed	342	−.03	.386	−0.10						
	Read the book before lecture	342							−.08	.020	−0.27
	Read the book after lecture	342							−.02	.626	−0.06
	Time studying	336	.02	.664	0.05						
	% of that time reviewing the book	342							−.09	.011	−0.29
	% of that time reviewing the lecture notes/slides	342							−.03	.416	−0.09
	% of that time reviewing the study guide	342							−.04	.322	−0.11
	Number of study sessions	342							.09	.017	0.27
	Number of timing strategies	342	.09	.015	0.28						

Note. Blanks = the strategy was not measured with the measurement type; p -values are provided without the correction; therefore, only those that are $p < .002$ are significant and are boldface

Discussion

The purpose of this work was to evaluate whether the ICAP framework could be extended the metacognitive study literature by examining students' self-reported study practices in an authentic classroom context. When integrating these two strands, two aspects of studying emerged: how one studied and time management. Interestingly, each strand contributed new strategies that the other had not included (e.g., metacognitive study strategies included time-management strategies, and ICAP included monitoring and regulation, among others). To assess these aspects of studying, students responded to closed-ended (Likert-scale, forced-choice) and open-ended questions before three noncumulative exams. Each type of measure had different affordances. The use of the open-ended question provided a census about which study strategies were most salient to students, as their responses were self-generated. In contrast, the use of the Likert-scale and forced-choice questions required students to make judgments about the described strategies. Across the measures, students reported using several types of strategies. In the next

few sections, we discuss the study and time-management strategies in terms of their reported use, their relation to each other (only for study strategies) and their relation to exam performance. Then we discuss the relations across the measures and some limitations and directions for future work.

Study strategies

The study strategies referred to the learning activities students engaged in to prepare for the exam. These included rewriting, highlighting, summarizing, generating examples, self-explaining, analogically comparing, quizzing/self-testing, metacognitive monitoring, and regulation. In general, the open-ended question revealed that students did not describe many study strategies. In fact, the most common strategy was monitoring, which only half the students reported using for any given exam. This strategy also had the largest variation in how often students reported using it in their open-ended responses. Otherwise, students were consistent in that they did not mention the other strategies often. In contrast, when prompted by the Likert-scale measures, students reported

using a moderate amount of constructive strategies. One reason for this discrepancy between student responses for the prompted and open-ended questions is that students might not be aware of the strategies that they use until prompted about them. Another reason for the discrepancy might be that when students are prompted with a strategy, it leads them to inflate their responses (see the Relation of the Measures section).

Although there was variation in the reports of student study strategies, we were able to categorize the self-reported strategies in alignment with the ICAP framework, representing two categories: active and constructive (Chi, 2009; Chi & Wylie, 2014). To further unpack the ICAP framework with self-reports, we also examined whether constructive and active strategies were related to each other. Some of the active and constructive strategies were positively related to each other, including creating examples and summarizing, and monitoring and highlighting. These results suggest that active and constructive strategies can work together in which the active strategies might help facilitate the use of constructive strategies. For example, in some of the responses, students would say they highlighted what they did not know, suggesting that when used together the active strategies can be used as a marker or referent for the constructive strategy to build upon. In this respect, it also supports the hierarchy assumption of the ICAP framework in which the higher categories subsume the processes of the lower.

We also examined the relation between the strategies within the same category: active or constructive. For the open-ended strategies, we found interesting patterns in which some of the active strategies were related to each other (rewriting and highlighting), but that these were not related to summarizing. Unlike the active strategies, each of the constructive strategies was positively related to another constructive strategy, with self-explanation being positively related to the most constructive strategies (creating examples, quizzing, comparison, and monitoring). These patterns were also supported in the Likert-scale measures in which all of the constructive strategies were positively related to each other. These findings suggest that the use of a constructive strategy tends to support the use of other constructive strategies, at least when students report using them. This result is also consistent with the ICAP hypothesis that constructive strategies share inference making processes and may be interrelated. For example, quizzing creates an opportunity to monitor one's knowledge and monitoring may facilitate self-explanation.

It is also important to note that there might be individual differences at play, such that students' motivational beliefs might be driving them to use constructive strategies. For example, if students had the goal to understand the material completely (a mastery-approach goal), we might predict that they will discover and use constructive strategies to accomplish that goal (Nokes-Malach & Mestre, 2013). They might

also be more likely to adopt other constructive strategies as opposed to active strategies that do not lead to conceptual understanding.

Moving beyond the relations between the strategies, we also examined whether these strategies were related to exam performance to further test the ICAP framework. The results were consistent the hypothesis that constructive strategies would have more positive relations to exam performance than the active strategies. Prior studies testing the hypothesis that constructive activities lead to better learning and performance than active have focused on observable learning activities. In the current work, we found similar results for students' self-reported study strategies. This adds to the growing evidence base for the positive relations between constructive study strategies and learning and performance outcomes across observational, experimental, and now student self-report data.

Time-management strategies

Additionally, we examined students' time-management strategies, which were unique to the metacognitive study strategy literature. In students' open-ended responses, time was not a salient aspect of their study descriptions. This finding suggests that students rarely think of time as being one way to study, which is consistent with Kornell and Bjork's (2007) interpretations that students do not think spacing one's study time is a strategy that helps memory. It could also be the case that answering the closed-questions about timing before the open-ended question resulted in some students not including those factors in their open-ended response because they thought that information was already accounted for in the prior responses. Alternatively, one might have expected that by bringing students' attention to those study features would have primed them to include those factors in their open-ended statements. Across all time-management variables, there was no relation to exam performance, which might be because time was not a salient factor in students' minds when it came to their studying practices and/or because time is difficult to estimate as evident with the large standard deviations for the amount of time studying and the percentage of time dedicated to different resources. More work is necessary to understand how students perceive their study time and the structure of their learning activities.

Relation of the measures: The importance of question type

The type of questions that assess student strategies is important to consider. Do similar measures align and do they have implications for how those responses relate to student learning? This work revealed that when framing the question to a specific exam, the two types of questions sometimes align, but that this depended on the construct. Students' open-ended

reports on self-explanation and analogical comparison—concrete and explicit strategies—were consistently related to their Likert-scale items. However, the open-ended monitoring statement was only related to the monitoring Likert-scale measure for the first and third exam. This relation disappeared at Exam 2. This result is not very surprising as metacognitive processes have a history of being difficult to adequately capture with Likert-scale measures (Winne & Perry, 2000; Zepeda, 2016). Another reason that this relation might have been weak is that, in hindsight, the Likert scales captured whether students thought that they were able to monitor, not specifically whether they did monitor.

Interestingly, the variation in measurement also revealed differences in their relation to exam performance. Self-explanation and analogical comparison were positively related to exam performance when they were measured by a Likert scale, but not by the open-ended question. One likely explanation is that students do not realize that self-explanation and analogical comparison are study strategies and thus are less likely to report these strategies, limiting the opportunity for them to be related to exam performance. Another possibility is that the Likert-scale measures prompt specific (and better) ways of using self-explanation and analogical comparison, which do not capture all the ways students engage and describe those strategies. For example, the self-explanation and analogical comparison items contained aspects of monitoring (which was positively related to exam performance for both types of measures), such that they required students to be able to know which parts of the material was difficult for them (e.g., “explain difficult concepts” and “If I don’t understand something”). Perhaps if the open-ended responses were coded in better alignment to the Likert-scale items, they would have had more converging relations.

The use of the open-ended question (although tedious to code) provides insight into the strategies that are at the forefront of students studying habits and reveal which strategies are overt and valued by students. A more specific frame to the question (a particular context vs. general) can also remove some of the difficulty in retrieving which strategies they used. Both of these adjustments might also help to alleviate the responses that could be inflated when general strategy prompts are provided. For comparison, in this work, students reported quizzing or testing themselves slightly more (an average of 19% across the exams) than the one prior study that used a general open-ended question (10.7%, Karpicke et al., 2009). These percentages are much lower in comparison to studies that specifically asked students if they generally used practice problems or tested themselves (71%, Hartwig & Dunlosky, 2012; 72%, Morehead et al., 2016).

There were also other differences between prior work and the research presented here. These differences included rewriting notes (averaged across exams 16% vs. 29.9%, Karpicke et al., 2009; 33%, Hartwig & Dunlosky, 2012; 33%,

Morehead et al., 2016), and highlighting (averaged across exams 13% vs. 6.2%, Karpicke et al., 2009; 22%, Hartwig & Dunlosky, 2012; 53%, Morehead et al., 2016). There was a similar response for creating examples (averaged across exams 4% vs. 4.5%; Karpicke et al., 2009). From these comparisons, it appears that the contextual framing of the question has implications for the frequency in which a strategy is reported, which can also have implications for the likelihood that it is related to performance. When assessing student study strategies, it may benefit researchers to further evaluate the nature of the questions they pose.

Limitations and future directions

Students who more frequently reported using (open-ended) or endorsing (Likert scale) constructive strategies were also students that performed well. Although these results are consistent with prior experimental work, we cannot conclude that they were causal in this study. It is possible that other unmeasured correlated variables were driving these effects (e.g., motivation; see Zepeda, Martin, & Butler, *in press*, for a commentary). Future work should further examine the relations between these study strategies, motivation, and learning outcomes. The current approach and methodology provide a set of measures that could be used in future work to examine whether a given instructional intervention changes students’ self-reported study strategies as well as their learning and performance outcomes.

The current work also focused on how students reported behaving outside of the class (their study strategies), but what was not clear was how the in-class demands interacted or affected how students behaved outside of the class. Although this work evaluated two semesters of the course, we did not examine whether the in-class demands (warm-up quizzes versus end-of-lecture quizzes) affected the types of strategies students reported. The types of activities and resources a course provides may result in differential learning outcomes. For instance, a classroom that provides resources that students can use to test themselves easily might result in more students reporting that they quizzed themselves.

Asking students to report on how they studied is a metacognitive process. Students who respond to questions about their studying have to both assess their awareness of their strategies and reflect on their study strategies. It is possible that asking students about their studying resulted in them engaging in more self-regulatory processes. For example, the questions might have prompted students to be more evaluative about the ways in which they studied, resulting in changes in their study strategies and subsequent learning. Future work could test whether responding to questions about studying throughout a course affects subsequent learning.

Another limitation of the current work is the retrospective nature of the self-reports. Whether or not the students were

accurate in their retrospective judgments of these strategies is still open for investigation. The current approach also does not capture the amount of time students spent using different strategies. An ecological momentary assessment would be a nice compliment to this work as it would also reveal when students use these study strategies and can be used to obtain a better estimate of how much time they spent using each type (Shiffman, Stone, & Hufford, 2008).

The type of knowledge that is covered in specific courses might also have implications for the strategies students report using and their relations to exam performance (Chi & Wylie, 2014; Wolters & Pintrich, 1998). For example, the course and exams in this study involved factual, conceptual, and applied knowledge in which the students had to know the topics covered, their conceptual underpinnings, and apply these concepts to different situations. Critically, the course went beyond covering only factual knowledge and did not require students to mathematically problem-solve. A course only emphasizing factual knowledge or requiring procedural knowledge with mathematical procedures may have different types of student-reported strategies and relations to exam performance. A productive line of research could examine the types of knowledge that self-reported strategies are related to across different domains. Courses that emphasize the link between procedural and conceptual knowledge or factual and conceptual knowledge and provide opportunities to apply their knowledge to new situations might have different relations between the study strategies and performance outcomes in comparison to courses that only emphasize the procedural, factual, or conceptual aspects.

Conclusions

Integrating the metacognitive study strategy and ICAP literatures provided several affordances in examining students' study practices such as categorizing study strategies based on theory (i.e., constructive and active categories via ICAP) and broadening the strategies to include additional ones. In support of the ICAP framework, we found that constructive strategies had more positive relations with each other and exam performance in comparison to active strategies. In particular, monitoring strongly predicted performance and was positively related to many of the other strategies. These results, along with the theoretical tie between monitoring and many of the constructive strategies, suggest that it may be a powerful construct to independently measure and incorporate as a separate category into the ICAP framework. Importantly, this study revealed that students' self-reported study strategies were predictive of learning and performance outcomes in theoretically consistent ways, providing support for researchers to use these self-reports as a measure in future study strategy experiments and interventions.

Acknowledgements This research was supported by the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition—Collaborative Grant No. 220020483 and Grant DUE—1524575 from the National Science Foundation. No endorsement should be inferred. We thank Drs. Tanner Wallace, Scott Fraundorf, and Chris Schunn for their feedback and guidance on the project. We also thank Mark Wertz, Emily Wenz, and Amanda Hopcroft for their help in coding the data.

Open practices statement The materials for the study are available upon request, the data for the study reported here are not available, and none of the study was preregistered.

References

- Aleven, V. A. W. M. M., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, 26(2), 147–179. doi:[https://doi.org/10.1016/S0364-0213\(02\)00061-7](https://doi.org/10.1016/S0364-0213(02)00061-7)
- Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, 48(2), 87–113. doi:<https://doi.org/10.1080/00461520.2013.775712>
- Bernaards, C. A., & Jennrich, R. I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, 65, 676–696.
- Bonett, D. G., & Wright, T. A. (2000). Sample size requirements for estimating Pearson, Kendall and Spearman correlations. *Psychometrika*, 65(1), 23–28. doi:<https://doi.org/10.1007/BF02294183>
- Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Mahwah, NJ: Erlbaum.
- Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105. doi:<https://doi.org/10.1111/j.1756-8765.2008.01005.x>
- Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, 18(3), 439–477. doi:https://doi.org/10.1207/s15516709cog1803_3
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. doi:<https://doi.org/10.1080/00461520.2014.965823>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences: Statistical power analysis for the behavioral sciences* (Vol. 2). Hillsdale, NJ: Erlbaum. doi:<https://doi.org/10.1234/12345678>
- Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The effect of instructional explanations on learning from scientific texts. *Journal of the Learning Sciences*, 6(4), 347–365. doi:https://doi.org/10.1207/s15327809jls0604_1
- Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, 21(4), 337–346. doi:<https://doi.org/10.1016/j.lindif.2011.03.002>
- 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(1), 4–58. doi:<https://doi.org/10.1177/1529100612453266>
- Dunn, K. E., Lo, W.-J., Mulvenon, S. W., & Sutcliffe, R. (2012). Revisiting the Motivated Strategies for Learning Questionnaire. *Educational and Psychological Measurement*, 72(2), 312–331. doi:<https://doi.org/10.1177/0013164411413461>
- Edwards, B. J., Williams, J. J., Gentner, D., & Lombrozo, T. (2019). Explanation recruits comparison in a category-learning task. *Cognition*, 185, 21–38. <https://doi.org/10.1016/j.cognition.2018.12.011>
- Ericsson, K. A., & Simon, H. A. (1980). Verbal protocols as data. *Psychological Review*, 87(3), 215–251.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109. doi:<https://doi.org/10.3102/00346543074001059>
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155–170. doi:[https://doi.org/10.1016/S0364-0213\(83\)80009-3](https://doi.org/10.1016/S0364-0213(83)80009-3)
- Gentner, D., Loewenstein, J., Thompson, L., & Forbus, K. D. (2009). Reviving inert knowledge: Analogical abstraction supports relational retrieval of past events. *Cognitive Science*, 33(8), 1343–1382. doi:<https://doi.org/10.1111/j.1551-6709.2009.01070.x>
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1–38. doi:[https://doi.org/10.1016/0010-0285\(83\)90002-6](https://doi.org/10.1016/0010-0285(83)90002-6)
- Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review*, 19(1), 126–134. doi:<https://doi.org/10.3758/s13423-011-0181-y>
- Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, 17(4), 471–479. doi:<https://doi.org/10.1080/09658210802647009>
- Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th). Upper Saddle River, NJ: Pearson Education International.
- Kim, S. (2015). ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, 22(6), 665–674. doi:<https://doi.org/10.5351/CSAM.2015.22.6.665>
- Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science*, 342(6161), 935–937. doi:<https://doi.org/10.1126/science.1238056>
- Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology*, 23(9), 1297–1317. doi:<https://doi.org/10.1002/acp.1537>
- Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, 14(2), 219–224. doi:<https://doi.org/10.3758/BF03194055>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. doi:<https://doi.org/10.1037/0022-3514.77.6.1121>
- McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition*, 39(3), 462–476. doi:<https://doi.org/10.3758/s13421-010-0035-2>
- Menekse, M., Stump, G. S., Krause, S., & Chi, M. T. H. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education*, 102(3), 346–374. doi:<https://doi.org/10.1002/jee.20021>
- Morehead, K., Rhodes, M. G., & Delozier, S. (2016). Instructor and student knowledge of study strategies. *Memory*, 24(2), 257–271. doi:<https://doi.org/10.1080/09658211.2014.1001992>
- Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning*, 1(2), 159–179. doi:<https://doi.org/10.1007/s10409-006-9595-6>
- Nokes, T. J., Hausmann, R. G. M., VanLehn, K., & Gershman, S. (2011). Testing the instructional fit hypothesis: The case of self-explanation prompts. *Instructional Science*, 39(5), 645–666. doi:<https://doi.org/10.1007/s11251-010-9151-4>
- Nokes-Malach, T. J., & Mestre, J. P. (2013). Toward a model of transfer as sense-making. *Educational Psychologist*, 48(3), 184–207. doi:<https://doi.org/10.1080/00461520.2013.807556>
- Nokes-Malach, T. J., VanLehn, K., Belenky, D. M., Lichtenstein, M., & Cox, G. (2013). Coordinating principles and examples through analogy and self-explanation. *European Journal of Psychology of Education*, 28, 1237–1263. doi:<https://doi.org/10.1007/s10212-012-0164-z>
- Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology*, 25(3), 241–286. doi:<https://doi.org/10.1006/ceps.2000.1040>
- O'Neil, H. F., Jr., & Abedi, J. (1996). Reliability and validity of a state metacognitive inventory: Potential for alternative assessment. *Journal of Educational Research*, 89(4), 234–245. <https://doi.org/10.1080/00220671.1996.9941208>
- Pashler, H., Zarow, G., & Triplett, B. (2003). Is temporal spacing of tests helpful even when it inflates error rates? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(6), 1051–1057. doi:<https://doi.org/10.1037/0278-7393.29.6.1051>
- Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor, MI: The University of Michigan. Retrieved from <https://files.eric.ed.gov/fulltext/ED338122.pdf>
- Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General*, 140(3), 283–302. doi:<https://doi.org/10.1037/a0023956>
- Rawson, K. A., Dunlosky, J., & Sciarrelli, S. M. (2013). The power of successive relearning: Improving performance on course exams and long-term retention. *Educational Psychology Review*, 25(4), 523–548. doi:<https://doi.org/10.1007/s10648-013-9240-4>
- Rawson, K. A., O'Neil, R., & Dunlosky, J. (2011). Accurate monitoring leads to effective control and greater learning of patient education materials. *Journal of Experimental Psychology: Applied*, 17(3), 288–302. <https://doi.org/10.1037/a0024749>
- Richey, J. E., & Nokes-Malach, T. J. (2015). Comparing four instructional techniques for promoting robust knowledge. *Educational Psychology Review*, 27(1), 181–218. doi:<https://doi.org/10.1007/s10648-014-9268-0>
- Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, 77(1), 1–15. doi:<https://doi.org/10.1111/j.1467-8624.2006.00852.x>
- Rittle-Johnson, B., & Star, J. R. (2011). The power of comparison in learning and instruction: Learning outcomes supported by different types of comparisons. In J. P. Mestre & B. H. Ross (Eds.), *Psychology of learning and motivation* (Vol. 55, pp. 199–225). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-0-12-387691-1.00007-7>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460–475. <https://doi.org/10.1006/ceps.1994.1033>
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4(1), 1–32. <https://doi.org/10.1146/annurev.clinpsy.3.022806.091415>

- Susser, J. A., & McCabe, J. (2013). From the lab to the dorm room: Metacognitive awareness and use of spaced study. *Instructional Science*, 41(2), 345–363. <https://doi.org/10.1007/s11251-012-9231-8>
- Tobias, S., & Everson, H. T. (2002). *Knowing what you know and what you don't: Further research on metacognitive knowledge monitoring* (College Board Research Report No. 2002-3). <https://eric.ed.gov/?id=ED562778>
- Vaughn, K. E., & Rawson, K. A. (2011). Diagnosing criterion-level effects on memory: What aspects of memory are enhanced by repeated retrieval? *Psychological Science*, 22(9), 1127–1131. doi:<https://doi.org/10.1177/0956797611417724>
- Walker, D. A. (2003). JMASM9: Converting Kendall's tau for correlational or meta-analytic analyses. *Journal of Modern Applied Statistical Methods*, 2(2), 525–530. doi:<https://doi.org/10.22237/jmasm/1067646360>
- Wang, M. T., & Degol, J. (2014). Staying engaged: Knowledge and research needs in student engagement. *Child Development Perspectives*, 8(3), 137–143. doi:<https://doi.org/10.1111/cdep.12073>
- Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). San Diego, CA: Academic Press.
- Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2012). How and when do students use flashcards? *Memory*, 20(6), 568–579. doi:<https://doi.org/10.1080/09658211.2012.687052>
- Wolters, C. A., & Pintrich, P. R. (1998). Contextual differences in student motivation and self-regulated learning in Mathematics, English, and Social Studies classrooms. *Instructional Science*, 26(August 1995), 27–47.
- Zepeda, C. D. (2016). *An investigation of three metacognitive measures (unpublished thesis)*. University of Pittsburgh, Pittsburgh, PA. Retrieved from <http://d-scholarship.pitt.edu/26563/>
- Zepeda, C. D., Martin, R. S., & Butler, A. C. (in press). Motivational strategies to engage learners in desirable difficulties. *Journal of Applied Research in Memory and Cognition*.
- Zhang, Z., & Yuan, K.-H. (2018). *Practical statistical power analysis using Webpower and R*. (Eds). Grandeur, IN: ISDSA Press.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183. doi:<https://doi.org/10.3102/0002831207312909>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.