American Educational Research Journal
June 2020, Vol. 57, No. 3, pp. 947–978
DOI: 10.3102/0002831219865221
Article reuse guidelines: sagepub.com/journals-permissions

Effects of a Science of Learning Course on College Students' Learning With a Computer

Jeffrey A. Greene
Nikki G. Lobczowski
Rebekah Freed
Brian M. Cartiff
University of North Carolina at Chapel Hill
Cynthia Demetriou
University of Arizona
A. T. Panter
University of North Carolina at Chapel Hill

First-year courses have been used to bolster college student success, but empirical evidence on their efficacy is mixed. We investigated whether a first-year science of learning course, focused on self-regulated learning, would benefit first-generation college students. We randomly assigned students to a treatment condition involving enrollment in the course, a comparison condition in which students had access to online course materials only, or a control condition. From this larger study, we recruited 43 students to participate in a laboratory task involving learning about the circulatory system with a computer. We found that treatment and comparison students experienced greater changes in conceptual knowledge than the control group, and we found differences in the enactment of monitoring and strategy use across conditions.

Keywords: computers, first-generation college students, science of learning, self-regulated learning

On average, individuals with a postsecondary degree earn more than their peers who lack one (Kena et al., 2015). Unfortunately, institutions of higher education in the United States continue to struggle to retain and graduate their students (Redford & Hoyer, 2017). Retention and graduation rates are particularly low for first-generation college students (FGCS; Lohfink & Paulsen, 2005; Redford & Hoyer, 2017), typically defined in the United States as those students whose parents have no postsecondary experience. Research has shown that FGCS, who on average comprise 33% of 4-year and community college populations, are less likely to be academically prepared for college and graduate at much lower rates than their peers whose parents have some postsecondary experience, despite FCGS' equivalent desire to succeed (Cahalan & Perna, 2015; X. Chen, 2005; Lauff & Ingels, 2013; Redford & Hoyer, 2017). This disparity between desire and result has led educators and researchers to explore multiple ways of supporting FGCS success.

College and university educators have implemented first-year courses in an attempt to bolster students' resilience, motivation, and knowledge, with presumed subsequent effects on retention and academic achievement (Hofer & Yu, 2003; Strayhorn, 2013; Young & Hopp, 2014). A recent meta-analysis of research on these courses revealed that, on average, they had very small effects on GPA (δ = .02) and retention (δ = .11) despite their

---

Jeffrey A. Greene is a professor of educational psychology and learning sciences in the School of Education at the University of North Carolina at Chapel Hill, 101A Peabody Hall, CB#3500, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599; e-mail: *jagreene@email.unc.edu*. His research interests include effective means of promoting learner achievement and digital literacy, particularly via helping students enact self-regulated learning and epistemic cognition. His work spans domains, including online learning, science, and history.

Nikki G. Lobczowski is a postdoctoral fellow in the Human-Computer Interaction Institute at Carnegie Mellon University. She completed her PhD in education at the University of North Carolina at Chapel Hill in the Learning Sciences and Psychology Studies Program. Her research interests include collaboration, social regulation of learning of cognition, motivation, emotions, and behaviors; designing technological interventions for K–12 STEM classrooms; and design-based research.

Rebekah Freed is a doctoral student in the Learning Sciences and Psychological Studies Program in the School of Education at the University of North Carolina at Chapel Hill. She holds a graduate degree in educational psychology and taught university courses for many years. Her research interests include how motivation and goal pursuit enhance self-regulated learning and conceptual understanding of social science.

Brian M. Cartiff is a doctoral student in the Learning Sciences and Psychological Studies Program in the School of Education at the University of North Carolina at Chapel Hill. He has degrees in chemistry and education and was a high school science teacher for more than 20 years. His research interests include epistemic cognition, critical thinking, and nature of science understanding, as well as how these factors influence public understanding of science.

Cynthia Demetriou is the Associate Vice Provost for Student Success and Retention Innovation at the University of Arizona where she is also a faculty affiliate in the Center for the Study of Higher Education and the Department of Educational Policy Studies.

A. T. Panter is the senior associate dean for undergraduate education for the College of Arts and Sciences and a professor of psychology and neuroscience in the L. L. Thurstone Psychometric Laboratory at University of North Carolina at Chapel Hill. She is past president of the American Psychological Association's Division on Quantitative and Qualitative Psychology. She develops instruments, research designs, program evaluations, and data-analytic strategies for applied research questions in higher education, personality, and health.

often-high financial and human resources costs (Permzadian & Credé, 2016). That meta-analysis revealed that courses focused on particular academic topics (e.g., history, science) were less effective than those focused on adjustment and orientation to higher education (e.g., review of campus resources and policies, instruction on study skills, time management, and learning strategies). The higher efficacy of courses focused on helping students "learn how to learn" (Hofer & Yu, 2003, p. 30) is not surprising when viewed through the lens of theory and research on self-regulated learning (SRL; Greene, 2018; Zimmerman, 2013). Self-regulated learners have indeed learned how to learn: They are able to pursue valued academic goals via planning, monitoring, controlling, and evaluating their cognition, motivation, behavior, and emotions, such as by using effective strategies and knowing how to self-motivate. The knowledge, skills, and dispositions to enact SRL are not intuitive or innate but can be acquired (Bjork, Dunlosky, & Kornell, 2013) and are predictive of academic performance (Dent & Koenka, 2016; Dignath & Büttner, 2008). Early research on SRL focused on its effects on achievement outcomes, but more recently researchers and educators have also become interested in SRL as a process (Azevedo, 2005; Greene et al., 2015; Schunk & Greene, 2018; Winne & Perry, 2000) and how students' enactment of SRL processes can be influenced (Dignath & Büttner, 2008; Graham, Harris, MacArthur, & Santangelo, 2018). Research findings suggest that the effects of first-year courses on retention and achievement may be significantly mediated by the degree to which they enhance students' ability to self-regulate their learning (Paris & Paris, 2001; Zimmerman, 2013); therefore, SRL may be a particularly important predictor of FGCS success.

In this study, which is part of a larger project on the short-term and long-term outcomes of a first-year science of learning course for FGCS (e.g., Hofer & Yu, 2003), we investigated whether this course had effects on a very specific, but common, aspect of the college student experience: learning with a computer-based learning environment (CBLE; Azevedo, 2005). We randomly assigned students to a treatment condition (i.e., enrollment in a face-to-face science of learning course), a comparison condition (i.e., access to readings from the course but no enrollment in the actual course), and a control condition (i.e., business as usual with no access to course or materials). Then, at the end of the semester, we recruited participants to engage in a laboratory study in which they were asked to think aloud (Ericsson & Simon, 1993; Greene, Deekens, Copeland, & Yu, 2018) as they used a computer to learn about the circulatory system. The think-aloud protocol (TAP) data allowed us to capture the kinds of SRL processing participants enacted, and we used pretest and posttest measures to assess changes in declarative and conceptual knowledge as a result of using the computer. We hypothesized that treatment participants would show greater changes in knowledge scores from pretest to posttest than from comparison or control participants. Also, we investigated how SRL processing differed across these groups and how this processing predicted changes in knowledge over the course of the task. These data afforded a window into both product (i.e., changes in knowledge) and process (i.e., TAP) effects of the science of learning course and shed light on how such courses change the ways in which FGCS enact the knowledge, skills, and dispositions necessary for success in higher education (Azevedo, 2005).

Literature Review

First-Generation College Students

Although researchers have defined FGCS status in a variety of ways (Demetriou, 2014), it is generally understood in the United States that FGCS are college students whose parents did not attend college or postsecondary education (Demetriou, Meece, Eaker-Rich, & Powell, 2017). Researchers and educators have endeavored to understand how to support FGCS and increase the likelihood of their college success (Perna & Thomas, 2006). Supporting FGCS benefits them during college as well as after they graduate (Garriott, Hudyma, Keene, & Santiago, 2015). Understanding the characteristics, challenges, and assets unique to FGCS is essential to determining how to support them effectively.

In terms of background, FGCS are more likely than non-FGCS, known as continuing-generation college students (CGCS), to be older, married, part-time students; live off campus; receive financial aid; and work full time while enrolled in school (Demetriou et al., 2017; Pike & Kuh, 2005). FGCS more often come from diverse sociocultural and socioeconomic backgrounds compared with CGCS (Demetriou, 2014). FGCS are more likely to be from an underrepresented ethnic background and to speak a home language other than English (Bui, 2002; Redford & Hoyer, 2017). FGCS often find differences between the sociocultural norms and values of their homes and their K-12 context, as well as their higher education context (Stephens, Markus, Fryberg, Johnson, & Covarrubias, 2015). These differences make it more difficult for FGCS to acquire college readiness knowledge and skills (e.g., SRL) during their K-12 education (Gamez-Vargas & Oliva, 2013; Stephens et al., 2015).

Studies suggest that FGCS are less academically prepared for college than CGCS (Choy, 2001; Redford & Hoyer, 2017). FGCS report feeling less prepared for college and fear failing in college more than CGCS (Pike & Kuh, 2005). They are more likely to have lower scores on the SAT, take fewer college credits, receive lower grades, need more remedial assistance, and withdraw from or repeat courses they attempt, as compared with CGCS (X. Chen, 2005). However, FGCS do not differ from CGCS in terms of their desire to succeed in college (Lauff & Ingels, 2013). Also, FGCS report knowing less about college social environments than their CGCS peers, which can lead to stress (Pike & Kuh, 2005).

Whereas many FGCS experience challenges in their transition to college, they are also likely to possess unique strengths. For example, in one study, FGCS of color from a low socioeconomic status demonstrated more resilience in the face of challenges than CGCS (Morales, 2014). As FGCS come from a more diverse socioeconomic and sociocultural background, they also demonstrate greater diversity in their goals (e.g., more group and community-related goals) that assist in completing the different types of coursework in college (Stephens et al., 2015). Depending on the context the students found themselves in, identifying as FGCS was also a source of strength for them in school (McKinley & Brayboy, 2004). In conclusion, FGCS who get into college certainly have the potential to succeed, but they are less likely to persist to graduation, on average, for a variety of reasons. It is important to provide environmental supports that help them succeed in college, so that they can persist to reach their educational goals. One support that universities can provide for FGCS is first-year courses.

First-Year Courses

Administrators and educators at many institutions of higher education have turned to first-year courses in their attempt to improve the academic performance and retention of their students. Retention benefits these institutions in terms of their academic mission of promoting student success, but also financially; replacing students who drop out costs more money than retaining them, and students doing better academically need fewer support services (Permzadian & Credé, 2016). Unfortunately, the empirical evidence regarding the efficacy of first-year courses, in terms of students' retention and academic performance, is mixed (Permzadian & Credé, 2016).

Four Main Types of First-Year Courses

Based on the literature, there are four main types of first-year courses: transition, academic themed, discipline themed, and remedial themed (Barefoot, 1992; Porter & Swing, 2006). Most institutions offer only one type of course, whereas some institutions provide a mixed format, in which students can choose from different types of seminar courses (Porter & Swing, 2006). Instructors leading first-year transition courses intend to help students acclimate to college life. Topics covered in these courses typically include study skills, connecting with faculty and staff, orientation to the campus and college life, and personal wellness (Rogerson & Poock, 2013). Findings on the effectiveness of this type of seminar are varied. Barton and Donahue (2009) found that enrollment in a first-year seminar transition course led to increases in GPA (Cohen's d = 0.19 as calculated based on reported t and df), but not retention, as compared with students who chose to take a shortened orientation course (i.e., either 1 week during the summer or a one-credit course during the fall semester). Cambridge-Williams, Winsler, Kitsantas, and Bernard (2013), on the other hand, found that students enrolled in transition courses had higher long-term academic retention (i.e., 4-year retention rate for treatment students was 75.0% compared with 59.9% for control), and higher graduation rates (i.e., 7-year graduation rate for treatment students was 68.7% compared with 55.9% for control), than students not enrolled in the courses. They also found that students self-reported higher short-term self-efficacy and SRL skills. Additionally, Miller and Lesik (2014) found evidence of higher retention for students in transition courses compared with students not enrolled in these courses, but only for retention during the students' first year (i.e., transition course students' odds of dropout were 40% compared with those of peers who did not participate in transition course) and graduation in their fourth year (i.e., transition course students were 1.9 times more likely to graduate).

A variety of instructors, including graduate students and full-time faculty from many disciplines, teach academic-themed first-year seminars. For these seminars, the instructors usually focus on an academic topic of their choice but are required to integrate common themes into their lessons (e.g., how to be reflective learners; Zerr & Bjerke, 2016). Research on this type of seminar is limited, but Zerr and Bjerke (2016) notably compared these academic-themed first-year courses with first-year transition courses and found no differences in student retention or GPA, but they did find higher student engagement. They also found that students rated these seminars negatively due to the amount of work required.

Discipline-themed first-year seminars offer students an introduction to a specific major or academic area (e.g., introduction to computer science), or preprofessional skills (e.g., public speaking or leadership; Black, Terry, & Buhler, 2016). Black et al. (2016) studied a variety of discipline-themed first-year courses and found that students enrolled in the specialized (e.g., introduction to computer science) and business-themed courses had statistically significantly higher retention rates from the fall to the spring semester than students enrolled in the generalized (e.g., elementary group dynamics) and English courses, as well as any courses tailored to transfer students.

Remedial-themed first-year courses focus on study skills and adjustment to college (e.g., time management, test preparation, and career planning). The main difference between remedial-themed and other types of seminars that also teach skills and adjustment techniques (e.g., transition seminars) is the type of students enrolled in the course. Namely, institution officials usually reserve remedial-themed first-year courses for at-risk, transfer, or struggling students (Forster, Swallow, Fodor, & Fousler, 1999). Researchers have reported that students enrolled in remedial courses exhibit higher grades than would be predicted based on their previous performance (Cone & Owens, 1991). These students also had higher retention rates (i.e., an increase in retention rate from 7% to 53%, Cone, 1991; Forster et al., 1999) and gains in study skills (Forster et al., 1999) compared with other at-risk cohorts that were not offered the courses.

Comparison of the Different Types of First-Year Seminars

Although many researchers have looked at these various types of courses separately, some have looked across the different types to determine which is the most effective. In a meta-analysis focused on first-year courses at the collegiate level, Hattie, Biggs, and Purdie (1996) investigated the effectiveness of teaching students learning or study skills, which were a central component to several different types of courses. They found that the overall effect of these courses at the university level (i.e., Cohen's $d = 0.27$) was below the average effect of other typical interventions in education. Despite more positive attitudes as a result of enrollment in these courses, students did not demonstrate substantive performance gains. The authors concluded that study skills training was relatively ineffective.

In another meta-analysis, Permzadian and Credé (2016) used categorizations from Barefoot (1992) to code seminars as extended orientation, academic, or hybrid. Unfortunately, they excluded seminars focused on study skills because Hattie et al. (1996) already had discussed the overall effectiveness of these types of courses. Overall, they found that first-year courses had almost no effect on first-year GPA ($\delta = .02$) and only a small effect on 1-year retention rates ($\delta = .11$). However, they found variance in efficacy. Using moderator analyses, they found the characteristics that had the largest effects on GPA were hybrid courses (i.e., a combination of extended orientation and academic content, $\delta = .11$), implementation at a 2-year institution ($\delta = .22$), studies published in a peer-reviewed publication ($\delta = .09$), and randomized ($\delta = .40$) or ex post facto with matching research designs ($\delta = .11$). For 1-year retention, the characteristics that produced the largest effects were extended orientation ($\delta = .12$), stand-alone courses (i.e., compared with those connected to other courses as part of a learning community, $\delta = .12$), courses taught by faculty or staff (i.e., compared with classes taught by graduate students, $\delta = .10$), and courses that targeted all first-year students ($\delta = .12$). Finally, they hypothesized that the low overall effect sizes may have been due to a lack of randomized experiments on these courses.

A Fifth Type of First-Year Seminar

Missing from these meta-analyses, however, is a fifth type of first-year seminar: learning to learn (i.e., science of learning). Even though these courses share some topics with transition courses (e.g., test-taking strategies, metacognitive skills), they are entrenched in educational and cognitive psychology research, including work on learning and memory (Hofer & Yu, 2003). These courses are typically inclusive of all students (i.e., not just remedial students). Although the literature on this type of course is limited, there are promising findings. For example, Tuckman and Kennedy (2011) found that students enrolled in learning to learn courses had higher GPAs (γ = .11), retention rates, and graduation rates (i.e., expected odds of graduation for students in the course were 1.69 times greater than control) than those first-year students not enrolled in the course. Hofer and Yu (2003) found that students in these courses reported lower test anxiety by the end of the semester and "significant increases in . . . memorization, elaboration, organization, deep processing, planning, and metacognition" (p. 31, Cohen's *d* ranged from 0.62 to 1.22). Finally, Wingate (2007) suggested that most study skills seminars do not support students holistically (i.e., being able to apply these skills to all courses) and that their focus should shift to learning how to learn (i.e., SRL), thereby changing students' "perceptions, learning habits, and epistemological beliefs" (p. 395).

Self-Regulated Learning

SRL refers to how "learners systematically activate and sustain their cognitions, motivations, behaviors, and affects" (Schunk & Greene, 2018, p. 1) toward attaining their personal learning goals. Quite a few distinct models of SRL have been proposed, but most of them have common assumptions: (1) learners are active participants in the construction of the knowledge; (2) learners have the abilities to observe and alter different aspects of their learning, as necessary; (3) learners enact the abilities in (2) based on their goals, criteria, or standards; and (4) that SRL processing serves as a mediator in the relationship between personal or contextual characteristics and learning outcomes (Pintrich, 2000). Most SRL models (e.g., Winne & Hadwin, 2008; Zimmerman, 2013) also include phases of SRL in which learners enact processes preceding learning (e.g., self-motivating, planning), during learning (e.g., monitoring progress against standards and toward goal achievement, changing plans or strategies as needed), and then after learning (e.g., reflecting, evaluating; Greene, 2018).

Research in SRL has developed greatly over the past 40 years (Winne, 2018), including investigations into the various kinds and aspects of cognition, metacognition, motivation, and strategy use that are critical for learning (Vandevelde, Van Keer, Schellings, & Van Hout-Wolters, 2015). These studies have been conducted with populations ranging from graduate students (Mullen, 2011; Whipp & Chiarelli, 2004), to undergraduate students (Kauffman, 2004; Moos & Azevedo, 2008), high school students (Winters & Azevedo, 2005), middle school students (Cleary & Kitsantas, 2017; Eom & Reiser, 2004; Kramarski & Mizrachi, 2006), and, more recently, elementary populations (Neitzel & Connor, 2017; Vandevelde et al., 2015), with the vast majority of these investigations revealing high correlations between effective SRL and learning performance. Also, research has shown strong relations between SRL and students' knowledge gains when using CBLEs to acquire conceptual understanding via text-based and multimodal information sources (Azevedo, 2005; Dent & Koenka, 2016; Greene et al., 2015; Winters, Greene, & Costich, 2008). However, much of this research has been correlational in nature, precluding claims of causality.

The vast majority of empirical research has been focused on the predictive validity of various SRL processes, either individually (e.g., practice testing; Adesope, Trevisan, & Sundararajan, 2017) or aggregated in some manner (Cleary & Kitsantas, 2017; Deekens, Greene, & Lobczowski, 2018). Less has been done to investigate the relationships between various SRL processes using objective measures, such as observation and performance (Ben-Eliyahu & Bernacki, 2015), though existent findings show that learners who accurately evaluate and calibrate their comprehension (Alexander, 2013) have greater tendencies to recognize ineffective learning strategies and switch to new ones (Binbasaran Tüysüzoğlu & Greene, 2015; Dunlosky & Thiede, 2013). Likewise, Deekens et al. (2018) found more frequent enactment of monitoring processes predicted more frequent use of deep-level strategies (i.e., strategies that foster elaboration and recall), which in turn predicted learning performance, above and beyond the effect of prior knowledge. Clearly, there is a need for further research on the role of conditional, contingent, and adaptive SRL processing (e.g., planning, monitoring, enacting of strategies, and evaluating progress and learning; Ben-Eliyahu & Bernacki, 2015) in learning how to help students enact those processes efficiently and effectively, and how changes in SRL processing relate to academic outcomes (Schunk & Greene, 2018).

Intervention research has revealed that SRL can be taught within courses and that students do seem to benefit from such instruction (e.g., P. Chen, Chavez, Ong, & Gunderson, 2017; Zepeda, Richey, Ronevich, & Nokes-Malach, 2015). However, there have not been systematic investigations of whether SRL training effects transfer to other courses or context. If one of the purported goals of SRL research is to help students truly self-regulate, then instruction on SRL should transfer beyond the context in which it was learned (Alexander, Dinsmore, Parkinson, & Winters, 2011). This is one of the goals of first-year learning to learn courses (Hofer & Yu, 2003).

Measuring Self-Regulated Learning

One limiting factor in the research on how to bolster SRL has been the challenge of effective and efficient measurement of SRL processing. In the past, researchers have measured SRL predominantly by self-report surveys and other instruments administered outside the actual context in which the SRL occurred (i.e., offline measures; Winne & Perry, 2000) such as structured interviews and teacher ratings. These types of measures have their advantages, including the ease with which they can be administered in large-scale testing (Schellings & Van Hout-Wolters, 2011) and their general feasibility, but their asynchronous administration requires the assumption that learners, or their teachers, can accurately recall relevant cognitive, metacognitive, motivational, behaviors, and affective processing. Furthermore, self-report measures present a closed set of response options, generated by researchers, that may not represent all the ways in which learners engage in SRL or present those ways in language that learners can understand. There is evidence that scores on self-report measures may not correspond to actual learner behavior (Veenman, 2007, 2011a, 2011b; Winne & Perry, 2000). In addition, self-report instruments, such as the ubiquitous Motivated Strategies for Learning Questionnaire (Pintrich, Smith, Garcia, & McKeachie, 1991), are too coarse grained (i.e., at a level that does not pick up on contextual details; Karabenick & Zusho, 2015), which has impelled investigators to use online (i.e., capturing SRL as it occurs) methodologies, including trace data (Perry & Winne, 2006), eye tracking (Scheiter, Schubert, & Schüler, 2018; Trevors, Feyzi-Behnagh, Azevedo, & Bouchet, 2016), and verbal report protocols (Azevedo & Cromley, 2004; Greene & Azevedo 2009; Greene, Deekens, et al., 2018).

Ericsson and Simon (1980, 1993) established that TAPs could help researchers gain insight into thinking without being too disruptive. These protocols involve learners verbalizing, but not explaining, all of their thoughts while working on a learning task (e.g., "I didn't understand what I just read, I'm going to read it again"). These data can reveal SRL processing (e.g., judgment of learning). Veenman, Elshout, and Groen (1993) found that TAPs might slow learners' performance in a learning activity, but they do not seem to alter regulatory processing. Subsequent research (e.g., Bannert & Mangelkamp, 2008) has supported these conclusions.

Perhaps more important, many researchers have successfully employed TAPs to reveal relations between online SRL processing and learning outcomes (e.g., Azevedo, Moos, Greene, Winters, & Cromley, 2008; Greene & Azevedo, 2007; Vandevelde et al., 2015). TAPs provide data that can capture SRL's dynamic, conditional, contingent, and contextual aspects (Greene, Deekens, et al., 2018; Ben-Eliyahu & Bernacki, 2015) that retrospective measures such as self-report instruments cannot elicit (Greene et al., 2015). TAPs' open-ended nature also provides richer data than self-report instruments because participants are able to verbalize freely, instead of being restricted to Likert-type items found in surveys (Greene, Robertson, & Costa, 2011). Such data allow students to share what they perceive they are actually doing and thinking, as opposed to being forced to choose among only those activities listed on a self-report measure.

Purpose of This Study

Research on academic performance suggests that the majority of college students and, in particular, many FGCS, would benefit from additional instruction in SRL (Greene, 2018; Zimmerman, 2013). First-year seminars can provide an extended, focused opportunity to directly instruct SRL and provide opportunities for practice with support and feedback. In this study, we randomly assigned FGCS to a treatment science of learning course with a strong focus on SRL, a comparison condition in which students had access to course materials but no actual instruction or support, or a business-as-usual control condition. Then, at the end of the semester, we recruited these college students to participate in a laboratory study of their ability to enact SRL while using a computer to learn. We chose this method of data collection to gather multimodal data on how FGCS students transferred their SRL skills to new learning contexts. We had one hypothesis and two research questions. First, we hypothesized that treatment participants would outperform participants in the other groups in terms of their acquisition of declarative and conceptual knowledge as a result of learning with the computer. Our research questions were "In what ways do participants in each group enact SRL processing differently?" and "How does SRL processing relate to conceptual knowledge at posttest?" We expected that participants who took the science of learning course would more frequently enact SRL processing, including more frequent monitoring and use of effective strategies, than students in the comparison or control conditions. We did not have expectations about differences in SRL processing between comparison and control conditions.

Method

Participants

During spring 2016, fall 2016, and spring 2017 semesters, we recruited FGCS at a highly selective university in the southeastern United States via email to participate in a study on the effects of a science of learning course on FGCS success. The students who agreed to participate in the study were randomly assigned into the experimental course condition to take the science of learning class, a comparison condition that had access to course materials online but did not enroll in the actual course, or a control condition where they did not take the course or receive access to any course materials. From this larger study including 137 students, we recruited 43 students to participate in the laboratory study described here. Of those 43 students who agreed to participate in the laboratory study, 20 participants were from the science of learning course condition, 11 participants were from the comparison condition, and 12 students were from the control condition. Recruitment for the laboratory portion of our study was a challenge, despite high participant incentives, requiring us to extend data collection across multiple semesters and sections of the science of learning course. Participants' ages ranged from 18 to 35 years, with a mean age of 20.72 years (SD = 3.14). There were 11 male and 32 female participants. Average reported GPA was 3.08 (SD = 0.46), with a range of 1.90 to 3.89. The most common majors were psychology (13), biology (8), and exercise and sport science (6).

Procedures

This science of learning course was designed as a learning to learn class (e.g., Hofer & Yu, 2003). The course was part of a U.S. Department of Education (2016) "First in the World" grant, called The Finish Line Project (FLP; https://www2.ed.gov/programs/fitw/index.html). The FLP focused on piloting well-researched programs and supports that increase FGCS success in college and progress to degree. As a part of the FLP grant, this study was advertised to FGCS via email. It was explained to students that if they agreed to participate in this study, they would be randomly placed into one of the three conditions, and that, as they participated in the study, they could be eligible to participate in a secondary research study, which would involve separate compensation. To ensure equitable benefits for all participants, those FGCS randomly assigned to comparison or control conditions were told that they would have priority if they wanted to take the course the following semester. For this laboratory study, students could only participate in their first semester in the larger study (e.g., a student who was in the comparison condition in spring 2016 and did not participate but then was placed in the treatment class in fall 2016 would be ineligible to participate in this study, because they had access to course materials for two semesters).

Toward the end of the semester in which they participated, students were invited to take part in a second laboratory portion of the research study. For this laboratory part of our research, we advertised for participants in each condition differently. In the treatment condition, we advertised by emailing the course students twice, and we advertised in class once. We advertised in the comparison and control conditions by emailing participants twice. We offered monetary compensation, in the form of a gift card redeemable at multiple businesses (e.g., Apple, Target) to incentivize participation in the laboratory study.

The Science of Learning Course

The science of learning course met twice a week for 1 hour and 15 minutes each. It was 3 credit hours, and there were between 13 and 20 students over multiple sections of the course and three semesters. In this course, students were expected to gain an understanding of the conceptual and empirical foundations of the science of learning, and they were also asked to apply this understanding to coursework, exams, and their own education. The goals for this course were for the students to be able to critically evaluate learning and education claims in the scholarly literature and media as well as be knowledgeable on how student motivation, deep learning strategies (Dinsmore, 2017), and self-regulation (Zimmerman, 2013) relate to academic success. Topics discussed in the course included academic SRL, internal and external factors of motivation, growth mind-set (i.e., implicit theories of intelligence; Dweck & Master, 2008), the information processing system, knowledge and expertise, intelligence, goal setting, emotions and cognition, technology and learning, sleep, exercise, and deep learning strategies, including spaced versus massed practice and elaboration. Specifically, we provided students with direct instruction on SRL and strategies and how they could use them to succeed in college. We used Zimmerman's (2013) model of SRL as a foundational aspect of the course, in terms of both providing students with an understanding of how to learn more effectively and also as an organizer of other topics in the course such as the roles of monitoring and strategy use in learning. In addition, we gave them opportunities in class to practice the various aspects of SRL during authentic learning activities and provided them with feedback.

We followed numerous procedures to ensure consistency in the course over the three semesters. First, we had the same two instructors teach the course. In the first semester, both instructors taught together (i.e., one instructor of record, one teaching assistant). For the next semester, the two instructors each taught a class separately. In the last semester, one of the instructors taught the course. Next, we created common assessments to use in each course across all three semesters, including weekly quizzes, five homework paper assignments, a midterm exam, and a final exam. Students completed the homework paper assignments independently, reflecting on how the material applied to their own personal and academic experiences (e.g., reflecting on their own SRL skills and identifying strengths and areas for improvement). Finally, the lesson plans were also similar across semesters. These lessons consisted of short lectures, class discussions, partner activities, small group discussions, videos and other media, and formative assessment questions using PowerPoint. Overall, our efforts to ensure consistency across semesters added to the quality and rigor of our research design.

Laboratory Study Procedures

To conduct our laboratory study of how students used a CBLE to acquire conceptual understanding, we used a procedure similar to the one used by Azevedo and colleagues (Azevedo, Gutherie, & Seibert, 2004; Azevedo, Johnson, & Burkett, 2015). Each learning session was conducted in a room with one participant and one researcher. First, the participants were informed of the length of time for the learning session and that they were able to opt out at any time without penalty. No participants chose to opt out. On agreeing to continue, the participant read and signed the laboratory study consent form that was approved by the university's institutional review board. Then, the participants had 15 minutes to complete the pretest. Participants were given the instructions to complete each page in the order provided, without skipping ahead or moving backward. The participants received no instructional material during the pretest and they were not told that the posttest would be identical to the pretest.

Next, on completion of the pretest, the participants received a tour of the relevant CBLE pages on heart, blood, and the circulatory system, within Microsoft Encarta (Microsoft Corporation, 2007). Then, participants were taught how to think aloud (Ericsson & Simon, 1993; Greene, Deekens, et al., 2018), including instructions to verbalize all thoughts and reading while interacting with the CBLE. Participants were able to practice thinking aloud within the CBLE using text irrelevant to the learning task, and researchers provided feedback on their verbalizations. Participants were told that they should verbalize but not explain their thinking, per Ericsson and Simon's (1993) protocol.

At this point, the participants were presented with the learning task in written form, which asked them to use the CBLE to learn about the circulatory system. The researcher read this task aloud to them, which included the following in bold: "Make sure you learn about the different parts and their purpose, how they work both individually and together, and how they support the human body." The learning task, which also reminded the participants to think aloud while learning, was posted next to the computer. Participants were given 30 minutes to learn in the CBLE without interruption. Participants were allowed to take notes, but were not required to do so, and informed that they would not be able to use them for the posttest.

The researcher audio- and videotaped the learning session, and we used screen capture software to record the computer screen. If the participants were silent for more than 2 seconds, the researcher prompted them to verbalize their thoughts by saying, "Keep talking, please." Such prompts occurred rarely and sporadically. Additionally, verbal time prompts were given to the participant at 20, 10, and 2 minutes left. After the 30 minutes, all recording devices were turned off, the CBLE was closed, and any notes taken by the participants were collected.

Last, participants were told that they had 15 minutes to complete the written posttest. They were not given access to the learning materials or their notes. On completion of the posttest, participants completed a demographic questionnaire. After the demographic questionnaire was completed, the researcher debriefed each participant about the study and asked each one of them to refrain from sharing the specific details of their participation in the study with other potential participants. In total, participant time in the laboratory study did not exceed 90 minutes.

Measures

Knowledge Measures

We collected pretest and posttest knowledge measures based on the work of Azevedo et al. (2004). These included matching (i.e., connecting

14 terms with their appropriate definition) and labeling (i.e., correctly identifying various parts of the heart) tasks, as well as a conceptual essay. For our declarative measures, two trained graduate students scored the matching and labeling portions of each participant's pretest and posttest for accuracy against an answer key. To score the conceptual knowledge measures, our mental model rubric included 12 possible scores that progressed from no understanding to complete understanding of the circulatory system. The 12 scores were (1) no understanding, (2) basic global concept, (3) basic global concept with purpose, (4) basic single-loop model, (5) single-loop model with purpose, (6) advanced single-loop model, (7) single-loop model with lungs, (8) advanced single-loop model with lungs, (9) double-loop concept, (10) basic double-loop model, (11) detailed double-loop model, and (12) advanced double-loop model. The 12 scores can be clustered into three categories (i.e., low, intermediate, and high conceptual understanding) based on understanding of the circulatory system as involving lungs (i.e., intermediate) and consisting of a double loop (i.e., high). Inter-rater reliability was calculated. The overall inter-rater reliability between scorers was 77.34%.

Think-Aloud Protocol Coding

We used TAPs to capture students' SRL. We transcribed these TAPs and coded them using a codebook and scheme from Azevedo et al. (2008), adapted for this study (see Online Supplementary File A in the online version of the journal). Within our coding scheme, 31 microlevel self-regulatory processes (e.g., taking notes, monitoring use of strategies) could be coded based on transcript content. Segments consisted of a word or a group of words that corresponded to one of the microlevel SRL processes. For example, if a participant stated, "I don't understand that," the statement would be coded as a Judgment of Understanding, with a negative valence (JOU−). If a segment was not codable or multiple microlevel codes could have reasonably been assigned to it, we coded the segment as No Code (NC) and ignored it during the analyses. Each microlevel code also fell within a macrolevel code category, such as planning or strategy use (Greene & Azevedo, 2009). We also aggregated certain microlevel codes into macrolevel deep-level strategy use and surface-level strategy use variables (Dinsmore, 2017; see Online Supplementary File B in the online version of the journal). These macrolevel variables were developed based on evidence in the extant research stating which high-utility strategies facilitate deep-level learning or surface-level learning of the material (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013) as well as the research conducted by Deekens et al. (2018). We chose two microlevel SRL codes, Feeling of Recognition, with a negative valence (FOR−), and JOU− to investigate based on Binbasaran Tüysüzoğlu and Greene (2015) and Ben-Eliyahu and Bernacki's (2015) work on the contingent nature of SRL. This work suggests that learners who are more accurate in assessing their understanding are also more likely to enact metacognitive control and enact effective SRL strategy use (Zimmerman, 2013).

Demographic Questionnaire

The demographic questionnaire included questions about participants' gender, age, and academic major. It asked participants to list the biology courses they had previous taken and designate whether or not the circulatory system was covered in each course listed. Last, the questionnaire asked participants to list relevant work experience related to health or medicine.

Coding Process for TAP Data

As mentioned previously, each participant transcription was coded twice, independently, by two members of the research team. After coding the transcripts independently, the researchers met to resolve any discrepancies in their coding. The first author was consulted if the two research team members could not come to a consensus on a given code. This type of coding is sometimes referred to as a "two-pass" coding. It is based on work conducted by Chi (1997) and is similar to procedures used by Herrenkohl and Cornelius (2013). Using a two-pass coding takes into account the complexity of coding and ensures that at least two researchers agree on a code for a given segment. Using this procedure ensures greater objectivity via consensus than methods where a single researcher codes a subset of the transcriptions, which is typical in studies that report inter-rater reliability. Additionally, previous research has shown that independent coders can achieve an acceptable level of reliability using this coding scheme (e.g., Greene & Azevedo, 2009).

It has been argued that reliability in coding methods may not be as important as predictive validity (Hammer & Berland, 2014). We assert that statistical calculations of inter-rater agreement are not particularly appropriate for this study. This is because two research team members coded each segment of every transcription independently, reconciled their coding, resolved disagreements, and came to a consensus through a process of consultation. Indeed, measures of reliability are more appropriately used in studies where raters code some subset of the data without verification from another researcher. This was not the procedure for this study. In fact, calculations of SRL TAP data coding agreements are rarely performed. This is due to the complexity of coding, as well as because researchers prioritize accuracy over inter-rater reliability. Though it is rare, when they are calculated, measures of SRL TAP coding agreements tend to be conducted on a very small subset of data (Bannert, Reimann, & Sonnenberg, 2014). In our case, our "two-pass" method, which ensures that two coders agree on every code assigned to every transcription, reflects our emphasis on predictive validity over inter-rater reliability. We privilege this, as demonstrated by our methods, despite its increase on resource demands.

Furthermore, our coding scheme has been used in numerous studies (e.g., Greene, Copeland, Deekens, & Yu, 2018; Greene, Costa, Robertson, Pan, & Deekens, 2010; Greene et al., 2015), with findings aligned with SRL theory, showing support for the predictive validity of the coding. In addition, other research teams have employed coding schemes that were derived from Azevedo and Cromley (2004), as we have here. These numerous studies published employing these coding schemes are more evidence that our coding scheme can be reliably implemented by other research teams (e.g., Greene et al., 2015; Johnson, Azevedo, & D'Mello, 2011; Moos, 2013). An additional source of evidence that these procedures can produce research and inferences with strong validity and reliability has been the recent proliferation of research with TAPs that do not only derive from this work but also utilize similar coding schemes and inter-rater procedures (e.g., De Backer, Van Keer, & Valcke, 2014; Dinsmore & Alexander, 2016).

Data Analysis

To analyze the knowledge measure data, first we established that pretest and posttest scores across semesters were not statistically significantly different from each other with an analysis of variance (all $p > .05$). Given these findings, we collapsed the data across semesters and conducted separate mixed analysis of covariance (ANCOVA) analyses (i.e., 2×3, pretest-posttest knowledge scores, and three conditions) for each set of knowledge measures: the two declarative knowledge measures (i.e., matching and labeling) and the conceptual knowledge measure (i.e., essay mental model), with students' number of previously-taken relevant courses as a covariate to account for prior experience in the academic domain. To analyze the SRL TAP data, we used measured variable path analysis to investigate how prior knowledge (i.e., pretest scores and number of previous courses taken) predicted monitoring (i.e., FOR− and JOU−) and how those monitoring codes predicted deep-level and surface-level learning strategies and subsequent posttest performance, per Deekens et al. (2018). Also, we investigated whether the frequency of monitoring and strategy use differed by condition.

We treated the knowledge measure data as continuous, normally distributed variables. For the TAP data, we summed the frequency of each participant's codes and then accumulated the total of each microlevel code into corresponding macrolevel aggregated codes (i.e., deep-level and surface-level strategy use). We treated our SRL TAP data as counts in all measured variable path models (Greene, Costa, & Dellinger, 2011).

Table 1
Descriptive Statistics for Pretest and Posttests

	M (*SD*)	Range	Skewness (*SE*)	Kurtosis (*SE*)
Previous courses	1.65 (1.41)	0–6	1.03 (0.36)	1.15 (0.71)
Pretest matching	8.12 (3.85)	1–13	−0.15 (0.36)	−1.23 (0.71)
Pretest labeling	2.30 (3.11)	0–12	1.51 (0.36)	1.66 (0.71)
Pretest essay	6.79 (3.14)	1–12	0.11 (0.36)	−0.55 (0.71)
Posttest matching	11.42 (2.39)	5–13	−1.41 (0.36)	0.92 (0.71)
Posttest labeling	7.56 (3.10)	0–14	−0.76 (0.36)	0.25 (0.71)
Posttest essay	9.02 (2.92)	1–12	−0.68 (0.36)	−0.10 (0.71)
Feeling of recognition minus	2.07 (2.53)	0–13	2.43 (0.36)	8.20 (0.71)
Judgment of understanding minus	1.56 (2.14)	0–8	1.50 (0.36)	1.55 (0.71)
Deep-level strategy use	9.51 (8.75)	0–42	1.74 (0.36)	3.94 (0.71)
Surface-level strategy use	18.30 (11.05)	1–49	1.06 (0.36)	1.11 (0.71)

Note. M = mean; *SD* = standard deviation; *SE* = standard error.

Results

Descriptive Statistics, Correlation Matrix, and Semester Analyses

Through descriptive statistics we found that, on average, the participants' mean declarative (i.e., matching and labeling) and conceptual (i.e., essay) knowledge scores increased from pretest to posttest (see Table 1). During this task, participants verbalized a lack of recognition more often than a lack of understanding and more frequently enacted surface-level strategies than deep-level strategies. From a correlation matrix of the variables (see Table 2), we found strong positive, statistically significant relationships among many of the knowledge measures. Notably, there were correlations between the pretest labeling scores and the number of previous courses taken with similar content, the frequency of negative judgments of understanding and the use of surface-level strategies, the frequency of deep-level strategy use and essay scores, the frequency of negative feelings of recognition and judgments of understanding, and the frequency of deep-level strategy use and negative feelings of recognition. We did not, however, find statistically significant relationships between the frequency of surface-level strategy use and scores on any of the knowledge measures. On the other hand, the frequency of deep-level strategy use was positively, statistically significantly correlated with our posttest conceptual knowledge measure (i.e., essay). Finally, the number of previous relevant courses taken did not correlate with two of our three pretest knowledge measures, suggesting that it might contribute useful additional information beyond our knowledge measures; therefore, we included it as a covariate in our analyses.

Table 2
Correlation Matrix

		1	2	3	4	5	6	7	8	9	10	11
1	Previous courses	—	0.24	.46**	.19	.26	.33*	−.22	−.11	−.23	.03	.05
2	Pretest matching		—	.59**	.61**	.68**	.61**	.30	−.13	−.29	.05	−.17
3	Pretest labeling			—	.54**	.38*	.61**	.25	−.17	−.25	.20	.01
4	Pretest essay				—	.51**	.62**	.36*	−.02	−.25	.08	.01
5	Posttest matching					—	.73**	.22	−.001	−.46**	.08	−.24
6	Posttest labeling						—	.34*	−.06	−.40**	.25	−.01
7	Posttest essay							—	.09	−.09	.31*	−.24
8	FOR−								—	.33*	.52**	.04
9	JOU−									—	.21	.40**
10	Deep-level strategy use										—	.25
11	Surface-level strategy use											—

Note. FOR = Feeling of Recognition; JOU = Judgment of Understanding. Correlations involving variables with count distributions (i.e., FOR−, JOU−) should be interpreted with caution given that they violate the assumption of a normal distribution.
$^{*}p < .05$. $^{**}p < .01$.

Knowledge Score Analysis

To investigate our hypothesis, we conducted three repeated-measures ANCOVA analyses (i.e., one for each knowledge measure: matching, labeling, and essay), with knowledge score as the within-subjects factor (i.e., pretest to posttest), condition as a between-subjects factor, and number of relevant previous courses as a covariate. For each analysis, Box's test of equality of covariance matrices was statistically nonsignificant; thus, we found no evidence of any violation of this assumption. In each analysis, there was a statistically significant (all ps $< .001$) and practically significant (all partial η^2 = .384–.649) within-subjects main effect, indicating that, on average, participants in each condition increased their declarative (i.e., matching and labeling) and conceptual (i.e., essay) knowledge over the course of the task. Previous courses did not have a statistically significant main effect in any ANCOVA, nor was the interaction of this covariate and the within-subjects factor statistically significant except for the conceptual knowledge analysis (p = .006, partial η^2 = .177). Likewise, there were no statistically significant interactions between the within-subjects factor and condition in the declarative knowledge ANCOVAs, but there was a statistically (p = .040) and practically (partial η^2 = .152) significant interaction for the conceptual knowledge ANCOVA.

Examination of adjusted means (see Figure 1) showed that, on average, participants in the control condition possessed unexpectedly high conceptual knowledge of the learning task at pretest, but they did not gain much conceptual knowledge as a result of learning with the computer. Notably, this was not a ceiling effect, as control participants' mean score on the essay posttest was 8.89, well below the maximum possible score of 12 on the measure. On the other hand, both treatment and comparison participants showed strong changes in knowledge from pretest to posttest, with treatment participants' slope greater than that of comparison participants. These differences in conceptual knowledge acquisition begged the question of why or how participants learned during the task, which was investigated using our SRL TAP process data.

Process Data Analysis

Given the lack of condition differences on the declarative knowledge measures, for the SRL processing analysis we focused solely on the conceptual knowledge measures. Our research question regarding SRL processing was "In what ways do participants in each group enact SRL processing differently and how does SRL processing relate to conceptual knowledge at posttest?" To test relations among knowledge and SRL process variables, we posited a path model similar to the one tested in Deekens et al. (2018). In this model, prior knowledge (i.e., essay pretest) and previous relevant coursework predicted the frequency of monitoring enacted during

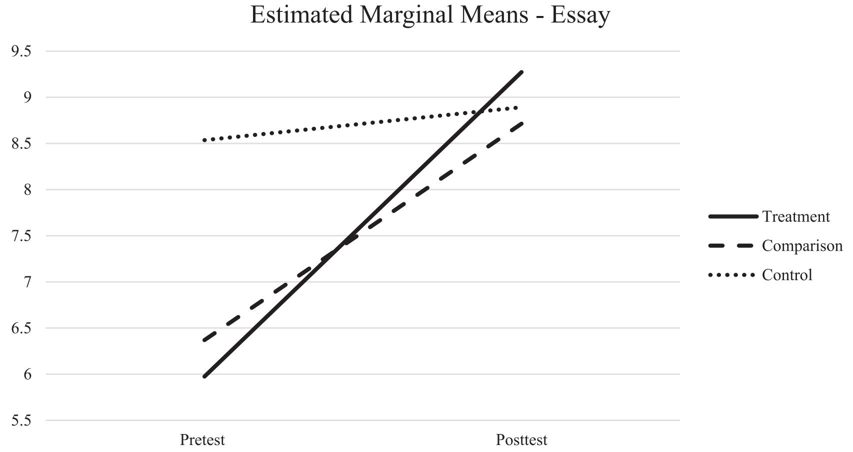

Figure 1. **Estimated marginal means for essay scores at pretest and posttest, by condition.**

learning (i.e., FOR− and JOU−), which in turn predicted the frequency of different kinds of strategies enacted (i.e., deep- and surface-level strategy use; Dinsmore, 2017), which then predicted performance on the knowledge essay posttest. The knowledge essay posttest was regressed on the knowledge pretest and previous coursework, as well. The condition to which participants were randomly assigned was posited to predict all monitoring, strategy use, and posttest variables, but in the final model, only those paths that were statistically significant were retained (see Figure 2). Finally, we tested paths from the essay pretest and previous coursework to all monitoring and strategy use variables and retained only those that were statistically significant. Based on examination of distributions and fit indices, we determined that the FOR− variable was best modeled as following a zero-inflated Poisson distribution and that the JOU-, deep-level strategy, and surface-level strategy variables were best modeled using a negative-binomial distribution (Greene, Costa, & Dellinger, 2011).

Our final model was estimated normally within Mplus 7.2 and had 31 free parameters, with a log-likelihood of −721.11. Neither data-model fit indices nor chi-square tests of data-model fit are available for path models with endogenous variables modeled with count distributions (i.e., zero-inflated Poisson, negative-binomial). The path coefficients illustrated in Figure 2 should be interpreted as the relationship between the criterion and the predictor above and beyond relationships between that criterion and other predictors.

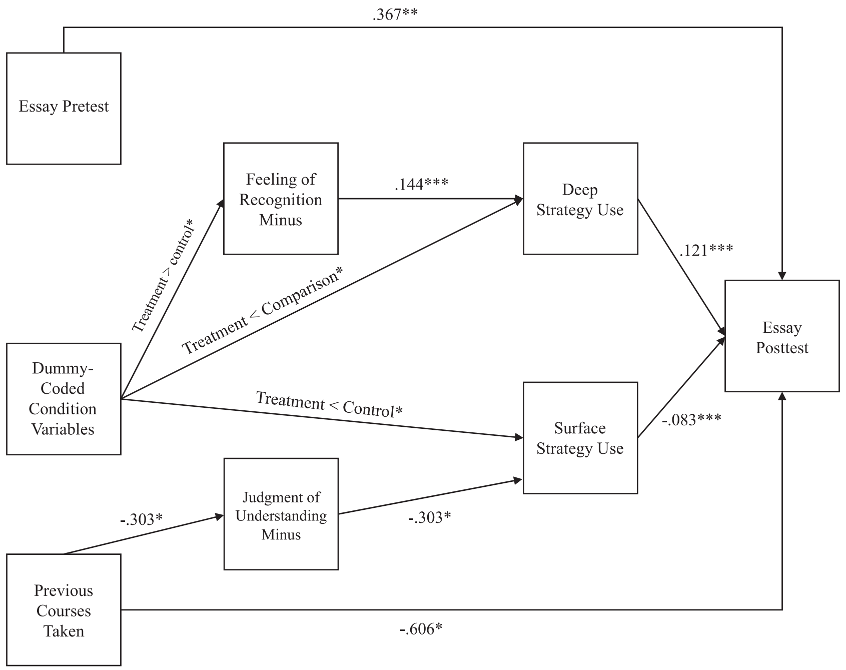

Figure 2. **Path analysis involving think-aloud protocol self-regulated learning (SRL) process data.**

Note. Statistically nonsignificant paths are not shown. "Dummy-coded condition variables" represent two dummy-coded condition variables. Paths representing variables regressed on dummy-coded condition variables illustrate the condition comparison that was statistically significant; all other comparisons were not statistically significant.
$*p < .05$. $**p < .01$. $***p < .001$.

As expected, pretest and posttest essay scores were positively related. More frequent use of deep-level strategies was positively related to posttest essay scores, whereas more frequent use of surface-level strategies was negatively related, again as would be expected given past research (e.g., Deekens et al., 2018; Dinsmore, 2017). Verbalizations coded as FOR− (e.g., "I do not think I have seen this before") were positively related to the frequency of deep-level strategy use, whereas verbalizations of JOU− (e.g., "I do not understand what I am reading") were negatively related to frequency of surface-level strategy use. As expected, participants who took more courses related to the biology content were less likely to verbalize that they did not understand something, but our measure of prior conceptual knowledge (i.e., essay pretest) was not statistically significantly related to any SRL monitoring or strategy use variables. The number of previous courses taken was negatively related to essay posttest scores, which was unexpected; this represents the relationship between such courses and posttest performance after accounting for prior knowledge. Finally, there were three relationships of note involving our conditions. Treatment participants more often verbalized FOR− than control participants, suggesting a difference in the use of monitoring between these groups. Treatment participants less often verbalized using deep-level strategies than the comparison participants, which was unexpected. Also, treatment participants less frequently reported using surface-level strategies than the control participants. Overall, these SRL process findings support the literature in terms of the posited relations among prior knowledge, monitoring, strategy use, and performance (Deekens et al., 2018; Dinsmore, 2017; Greene, 2018; Moos & Azevedo, 2008) and provide support for some but not all of our expected effects of our treatment on SRL processing.

Discussion

In the United States, low retention and graduation rates, particularly among FGCS, have led university administrators and researchers to turn to first-year courses in an attempt to improve FGCS' skills and increase their likelihood of success (Redford & Hoyer, 2017; Young & Hopp, 2014). Unfortunately, the efficacy of these seminars remains unclear (Permzadian & Credé, 2016). Existing empirical evidence favors first-year courses focused on teaching students the science of learning, in particular SRL knowledge, skills, and dispositions (Hofer & Yu, 2003; Zimmerman, 2013). Therefore, in this study, we randomly assigned FGCS to a science of learning course, a comparison condition with access to the materials from the course but no instruction, or a control business-as-usual condition. Then, we recruited 43 of these FGCS to participate in a laboratory study of how they used a computer to learn about the circulatory system. We found that participants, on average, increased both their declarative and conceptual knowledge over the course of the learning task. However, treatment and comparison condition participants experienced greater increases in conceptual knowledge scores from pretest to posttest than control students, with the treatment participants achieving the largest increase. We had hypothesized that the treatment condition participants would outperform the other groups; therefore, the comparison groups' performance was unexpected.

Our research question involved how SRL processing related to performance and what differences existed in that processing across conditions. The frequency of deep-level strategy use positively predicted conceptual knowledge performance, whereas more frequent use of surface-level strategies was negatively related to such performance, as found in other research (Deekens et al., 2018; Dinsmore, 2017). Participants who more frequently verbalized a feeling of not recognizing something also more frequently enacted deep-level strategies, suggesting that participants were using monitoring information to control learning and calibrate their learning strategies to their needs (Alexander, 2013; Zimmerman, 2013). The more often a participant verbalized a lack of understanding, the less often they stated that they were using surface-level strategies, which further supports the theoretical connection between monitoring and control (Ben-Eliyahu & Bernacki, 2015). These findings, which were common across conditions, support models of SRL in that they show the sequential and conditional nature of monitoring, strategy use, and learning (Ben-Eliyahu & Bernacki, 2015; Binbasaran Tüysüzoğlu & Greene, 2015; Zimmerman, 2013).

Our science of learning course included a focus on using deep- as opposed to surface-level strategies (Dinsmore, 2017), and we found treatment participants did this via the path through monitoring, supporting the role of SRL as a mediator of the relationship between individual differences such as prior knowledge and learning performance (Zimmerman, 2013). This suggests that the course was effective in helping FGCS acquire monitoring and strategy use knowledge and skills, two phenomena thought to be important across many contexts (Alexander et al., 2011; Greene, 2018). On the other hand, comparison participants more frequently verbalized using deep-level strategies than treatment participants. These analyses shed light on why treatment and comparison participants acquired more conceptual knowledge over the course of the learning task than their control condition peers: They more frequently enacted the kinds of monitoring and strategy use associated with learning (Dinsmore, 2017; Dunlosky et al., 2013).

To our knowledge, this is the first study to utilize random assignment as well as both product and process data (Azevedo, 2005) to investigate the effects of a first-year science of learning seminar on FGCS. As such, it makes a significant contribution to the literature regarding how to improve the efficacy of first-year courses by focusing them on SRL and strategy use (Permzadian & Credé, 2016). Such seminars provide FGCS with the knowledge, skills, and dispositions necessary for success in college, and in the case of this study, those necessary for enacting what has become a key component of college success: learning with a computer (Greene, Copeland, et al., 2018). Such transfer, from the classroom to the laboratory and computer context, is promising and deserves further research (Alexander et al., 2011). Researchers should also compare learning to learn courses with other first-year initiatives, with populations including both FGCS as well as other students (e.g., general, at-risk, and transfer students). Additionally, researchers can extend our study to capture the effect of courses like ours on learning within the course itself by analyzing in-vivo classroom academic outcome or trace data students generate when interacting with technology (e.g., learning management systems; Bernacki, 2018).

Limitations

There are a few limitations regarding the design of our study. First, we had issues with recruitment, despite generous incentives. This required us to incorporate data collection from multiple semesters, which despite no systematic differences in course instruction of student pretest performance in our laboratory study, may have introduced some confounding effects. Our difficulties with recruitment also resulted in a less than optimal sample size and concerns about selection effects. Our small sample size limited us to being able to only investigate a subset of the possible relations among SRL and learning. Finally, the nature of the laboratory setting in which we studied the participants' learning processes limits the external validity of our findings, thus restricting our ability to generalize to different learning environments.

Future Directions for Practice

As institutions of higher education expand their study of distal outcomes (e.g., retention, graduation) to include proximal ones such as student satisfaction and lifelong success, there is a growing need for rigorous research on not only what works but also why, as required by the What Works in Education Clearinghouse and other education organizations (Honig, 2009; Kinzie & Kuh, 2016). Our findings, based on a randomized control trial, support the development of first-year courses focused on the science of learning with an emphasis on explicit instruction of SRL as well as frequent opportunities for practice and feedback. In particular, the first specific suggestion for developers of first-year courses, derived from our findings, is that SRL should be taught as a set of related processes, rather than separate components, with ample opportunities for practicing monitoring and control in authentic contexts so that students learn how calibration informs effective learning strategy use (Paris & Paris, 2001; Weinstein & Acee, 2013). Second, many students reported a lack of knowledge regarding the research on effective strategy use (e.g., Dunlosky et al., 2013); therefore, explicit instruction and practice using deep strategies is also likely a key component of a successful learning to learn course (Zepeda et al., 2015). Finally, we suggest that course developers include frequent opportunities for students to reflect on their academic work and how SRL could be used to strengthen it. Course evaluation and anecdotal feedback indicated that students found such activities very helpful. All these suggestions should be informed by other research on first-year seminars, such as Permzadian and Credé's (2016) findings that the most efficacious courses combined extended orientation and academic content; were not connected to other courses as a part of a learning community; were taught by faculty or staff, not graduate students; and were targeted for all first-year students. Finally, as Wingate (2007) suggested, instructors should explicitly endorse students' generalization of deep learning strategies beyond a single course to promote true self-regulation of learning throughout their college career and life.

Conclusion

In sum, we found evidence that a science of learning course can help FGCS acquire conceptual knowledge using a computer via more effective SRL processing. This evidence of transfer from the classroom to the laboratory task is compelling and makes a strong case for continued development and refinement of first-year seminar learning to learn courses as a way to bolster academic performance and retention. More research is needed to determine whether such first-year seminars lead to transfer of SRL knowledge, skills, and dispositions to other college courses and outcomes. Our study provides a model for doing such work utilizing online measures of SRL and suggests the need for future research in authentic contexts.

Notes

Supplemental material is available for this article in the online version of the journal.

This research is part of the Finish Line Project (P116F140018; Panter, PI; Demetriou, Executive Director), which is funded by the U.S. Department of Education's "First in the World" grant program. The opinions expressed are those of the authors and do not represent the views of the U.S. Department of Education.

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, *87*, 659–701.

Alexander, P. A. (2013). Calibration: What is it and why it matters? An introduction to the special issue on calibrating calibration. *Learning and Instruction*, *24*, 1–3. doi:10.1016/j.learninstruc.2012.10.003

Alexander, P. A., Dinsmore, D. L., Parkinson, M. M., & Winters, F. I. (2011). Self-regulated learning in academic domains. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 393–407). New York, NY: Routledge.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, *40*, 199–209.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, *96*, 523–535.

Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research*, *30*, 87–111.

Azevedo, R., Johnson, A., & Burkett, C. (2015, July). *Does training of cognitive and metacognitive regulatory processes enhance learning and deployment of processes with hypermedia?* Paper presented at the Annual Meeting of the Cognitive Science Society, Pasadena, CA.

Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. C. (2008). Why is externally regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research & Development*, *56*, 45–72.

Bannert, M., & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted: Does the verbalization method affect learning? *Metacognition and Learning*, *3*, 39–58.

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning*, *9*, 161–185.

Barefoot, B. O. (1992). *Helping first-year college students climb the academic ladder: Report of a national survey of freshman seminar programming in American higher education* (Doctoral dissertation, College of William and Mary). Retrieved from https://scholarworks.wm.edu/cgi/viewcontent.cgi?article=1790&context=etd

Barton, A., & Donahue, C. (2009). Multiple assessments of a first-year seminar pilot. *JGE: The Journal of General Education, 58*, 259–278.

Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: A focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, *10*, 1–13.

Bernacki, M. L. (2018). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 370–387). New York, NY: Routledge.

Binbasaran Tüysüzoğlu, B., & Greene, J. A. (2015). An investigation of the role of contingent metacognitive behavior in self-regulated learning. *Metacognition & Learning*, *10*, 77–98.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417–444.

Black, A., Terry, N., & Buhler, T. (2016). The impact of specialized courses on student retention as part of the freshman experience. *Allied Academies International Conference: Academy of Educational Leadership*, *20*, 85–92.

Bui, K. V. T. (2002). First-generation college students at a four-year university: Background characteristics, reasons for pursuing higher education, and first-year experience. *College Student Journal*, *36*, 3–11.

Cahalan, M., & Perna, L. (2015). *Indicators of higher education equity in the United States: 45 year trend report*. Washington, DC: The Pell Institute.

Cambridge-Williams, T., Winsler, A., Kitsantas, A., & Bernard, E. (2013). University 100 orientation courses and living-learning communities boost academic retention and graduation via enhanced self-efficacy and self-regulated learning. *Journal of College Student Retention: Research, Theory & Practice*, *15*, 243–268.

Chen, P., Chavez, O., Ong, D. C., & Gunderson, B. (2017). Strategic resource use for learning: A self-administered intervention that guides self-reflection on effective resource use enhances academic performance. *Psychological Science*, *18*, 774–785.

Chen, X. (2005). *First generation students in postsecondary education: A look at their college transcripts* (NCES 2005-171). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, *6*, 271–315.

Choy, S. (2001). *Students whose parents did not go to college: Postsecondary access, persistence, and attainment* (NCES 2001-126). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Cleary, T. J., & Kitsantas, A. (2017). Motivation and self-regulated learning influences on middle school mathematics achievement. *School Psychology Review*, *46*, 88–107.

Cone, A. L. (1991). Sophomore academic retention associated with a freshman study skills and college adjustment course. *Psychological Reports*, *69*, 312–314.

Cone, A. L., & Owens, S. K. (1991). Academic and locus of control enhancement in a freshman study skills and college adjustment course. *Psychological Reports, 68*(3 Suppl.), 1211–1217.

De Backer, L., Van Keer, H., & Valcke, M. (2014). Promoting university students' metacognitive regulation through peer learning: The potential of reciprocal peer tutoring. *Higher Education*, *70*, 469–486.

Deekens, V. M., Greene, J. A., & Lobczowski, N. G. (2018). Monitoring and depth of strategy use in computer-based learning environments for science and history. *British Journal of Educational Psychology*, *88*, 63–79.

Demetriou, C. (2014). *Reflections at the finish line: The activities, roles, and relationships of college success for first-generation college students*. (Doctoral dissertation). Retrieved from https://cdr.lib.unc.edu/concern/dissertations/cj82k8251

Demetriou, C., Meece, J., Eaker-Rich, D., & Powell, C. (2017). The activities, roles, and relationships of successful first-generation college students. *The Journal of College Student Development*, *58*, 19–36.

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*, 425–474.

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students: A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*, 231–264.

Dinsmore, D. L. (2017). *Strategic processing in education*. New York, NY: Routledge.

Dinsmore, D. L., & Alexander, P. A. (2016). A multidimensional investigation of deep-level and surface-level processing. *Journal of Experimental Education*, *84*, 213–244.

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.

Dunlosky, J., & Thiede, K. W. (2013). Four cornerstones of calibration research: Why understanding students' judgments can improve their achievement. *Learning and Instruction*, *24*, 58–61.

Dweck, C. S., & Master, A. (2008). Self-theories motivate self-regulated learning. In D. Schunk & B. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 31–51). Mahwah, NJ: Lawrence Erlbaum.

Eom, W., & Reiser, R. A. (2000). The effects of self-regulation and instructional control on performance and motivation in computer-based instruction. *International Journal of Instructional Media*, *27*, 247–260.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 215–251.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge: MIT Press.

Forster, B., Swallow, C., Fodor, J. H., & Fousler, J. E. (1999). Effects of a college study skills course on at-risk first-year students. *NASPA Journal*, *36*, 120–132.

Gamez-Vargas, J., & Oliva, M. (2013). Adult guidance for college: Rethinking educational practice to foster socially-just college success for all. *Journal of College Admission*, *221*, 60–68.

Garriott, P. O., Hudyma, A., Keene, C., & Santiago, D. (2015). Social cognitive predictors of first- and non-first-generation college students' academic & life satisfaction. *Journal of Counseling Psychology*, *62*, 253–263.

Graham, S., Harris, K. R., MacArthur, C., & Santangelo, T. (2018). Self-regulation and writing. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 138–152). New York, NY: Routledge.

Greene, J. A. (2018). *Self-regulation in education*. New York, NY: Routledge.

Greene, J. A., & Azevedo, R. (2007). Adolescents' use of self-regulatory processes and their relation to qualitative mental model shifts while using hypermedia. *Journal of Educational Computing Research*, *36*, 125–148.

Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of sophisticated mental models. *Contemporary Educational Psychology*, *34*, 18–29.

Greene, J. A., Bolick, C. M., Jackson, W. P., Caprino, A. M., Oswald, C., & McVea, M. (2015). Domain-specificity of self-regulated learning processing in science and history. *Contemporary Educational Psychology*, *42*, 111–128.

Greene, J. A., Copeland, D. Z., Deekens, V. M., & Yu, S. (2018). Beyond knowledge: Examining digital literacy's role in the acquisition of understanding in science. *Computers & Education*, *117*, 141–159.

Greene, J. A., Costa, L.-J., & Dellinger, K. (2011). Analysis of self-regulated learning processing using statistical models for count data. *Metacognition & Learning*, *6*, 275–301.

Greene, J. A., Costa, L.-J., Robertson, J., Pan, Y., & Deekens, V. (2010). Exploring relations among college students' prior knowledge, implicit theories of intelligence, and self-regulated learning in a hypermedia environment. *Computers & Education*, *55*, 1027–1043.

Greene, J. A., Deekens, V. M., Copeland, D. Z., & Yu, S. (2018). Capturing and modeling self-regulated learning using think-aloud protocols. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 323–337). New York, NY: Routledge.

Greene, J. A., Robertson, J., & Costa, L.-J. C. (2011). Assessing self-regulated learning using think-aloud protocol methods. In B. J. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 313–328). New York, NY: Routledge.

Hammer, D., & Berland, L. K. (2014). Confusing claims for data: A critique of common practices for presenting qualitative research on learning. *Journal of the Learning Sciences*, *23*, 37–46.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, *66*, 99–136.

Herrenkohl, L. R., & Cornelius, L. (2013). Investigating elementary students' scientific and historical argumentation. *Journal of the Learning Sciences*, *22*, 413–461.

Hofer, B. K., & Yu, S. L. (2003). Teaching self-regulated learning through a "learning to learn" course. *Teaching of Psychology*, *30*, 30–33.

Honig, M. I. (2009). What works in defining "what works" in educational improvement: Lessons from education policy implementation research, directions for future research. In G. Sykes, B. Schneider, & D. N Plank (Eds.), *Handbook of education policy research* (pp. 333–347). New York, NY: Routledge.

Johnson, A., Azevedo, R., & D'Mello, S. (2011). The temporal and dynamic nature of self-regulatory processes during independent and externally assisted hypermedia learning. *Cognition and Instruction*, *29*, 471–504.

Karabenick, S. A., & Zusho, A. (2015). Examining approaches to research on self-regulated learning: Conceptual and methodological considerations. *Metacognition and Learning*, *10*, 151–163.

Kauffman, D. F. (2004). Self-regulated learning in web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing, and motivational beliefs. *Journal of Educational Computing Research*, *30*, 139–161.

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., Wilkinson-Flicker, S., Barmer, A., & Dunlop Velez, E. (2015). *The condition of education 2015* (NCES 2015-144). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Kinzie, J., & Kuh, G. (2016, November). *Report prepared for the Lumina Foundation: Review of student success frameworks to mobilize higher education*. Bloomington: Indiana University, Center for Postsecondary Research.

Kramarski, B., & Mizrachi, N. (2006). Online discussion and self-regulated learning: Effects of instructional methods on mathematical literacy. *Journal of Educational Research*, *99*, 218–230.

Lauff, E., & Ingels, S. J. (2013). *Education longitudinal study of 2002 (ELS: 2002): A first look at 2002 high school sophomores 10 years later* (NCES 2014-363). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Lohfink, M. M., & Paulsen, M. M. (2005). Comparing the determinants of persistence for first-generation and continuing-generation students. *Journal of College Student Development 46*, 409–428.

McKinley, B., & Brayboy, J. (2004). Hiding in the ivy: American Indian students and visibility in elite educational settings. *Harvard Educational Review*, *74*, 125–152.

Microsoft Corporation. (2007). Encarta premium. [Computer software]. Redmond, WA: Microsoft.

Miller, J. W., & Lesik, S. S. (2014). College persistence over time and participation in a first-year seminar. *Journal of College Student Retention: Research, Theory & Practice*, *16*, 373–390.

Moos, D. C. (2013). Examining hypermedia learning: The role of cognitive load and self-regulated learning. *Journal of Educational Multimedia and Hypermedia*, *22*, 39–61.

Moos, D. C., & Azevedo, R. (2008). Self-regulated learning with hypermedia: The role of prior domain knowledge. *Contemporary Educational Psychology*, *33*, 270–298.

Morales, E. E. (2014). Learning from success: How original research on academic resilience informs what college faculty can do to increase the retention of low socioeconomic status students. *International Journal of Higher Education*, *3*, 92–102.

Mullen, C. A. (2011). Facilitating self-regulated learning using mentoring approaches with doctoral students. In B. J. Zimmerman, D. H. Schunk, B. J. Zimmerman, & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 137–152). New York, NY: Routledge/Taylor & Francis Group.

Neitzel, C., & Connor, L. (2017). Messages from the milieu: Classroom instruction and context influences on elementary school students' self-regulated learning behaviors. *Journal of Research in Childhood Education*, *31*, 548–560.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, *36*, 89–101.

Permzadian, V., & Credé, M. (2016). Do first-year seminars improve college grades and retention? A quantitative review of their overall effectiveness and an examination of moderators of effectiveness. *Review of Educational Research*, *86*, 277–316.

Perna, L. W., & Thomas, S. L. (2006). *Commissioned report for the national symposium on postsecondary student success: Spearheading a dialog on student-success, a framework for reducing the college success gap and promoting success for all*. Washington, DC: National Postsecondary Education Cooperative.

Perry, N. E., & Winne, P. H. (2006). Learning from learning kits: gStudy traces of students' self-regulated engagements with computerized content. *Educational Psychology Review*, *18*, 211–228.

Pike, G. R., & Kuh, G. D. (2005). First- and second-generation college students: A comparison of their engagement and intellectual development. *The Journal of Higher Education*, *76*, 276–300.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic Press.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.

Porter, S. R., & Swing, R. L. (2006). Understanding how first-year seminars affect persistence. *Research in Higher Education*, *47*, 89–109.

Redford, J., & Hoyer, K. M. (2017). *First-generation and continuing-generation college students: A comparison of high school and postsecondary experiences*. Washington, DC: Institute of Education Sciences, National Center for Education Statistics.

Rogerson, C. L., & Poock, M. C. (2013). Differences in populating first year seminars and the impact on retention and course effectiveness. *Journal of College Student Retention: Research, Theory & Practice*, *15*, 157–172.

Scheiter, K., Schubert, C., & Schüler, A. (2018). Self-regulated learning from illustrated text: Eye movement modelling to support use and regulation of cognitive processes during learning from multimedia. *British Journal of Educational Psychology*, *88*, 80–94.

Schellings, G., & Van Hout-Wolters, B. H. (2011). Measuring strategy use with self-report instruments: Theoretical and empirical considerations. *Metacognition and Learning*, *6*, 83–90.

Schunk, D. H., & Greene, J. A. (2018). Historical, contemporary, and future perspectives on self-regulated learning and performance. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 1–15). New York, NY: Routledge.

Stephens, N. M., Markus, H. R., Fryberg, S. A., Johnson, C. S., & Covarrubias, R. (2015). Unseen disadvantage: How American universities' focus on independence undermines the academic performance of first generation college students. *Journal of Personality and Social Psychology*, *21*, 420–429.

Strayhorn, T. L. (2013). Academic achievement: A higher education perspective. In J. Hattie & E. M. Anderman (Eds.), *International guide to student achievement* (pp. 16–18). New York, NY: Routledge.

Trevors, G., Feyzi-Behnagh, R., Azevedo, R., & Bouchet, F. (2016). Self-regulated learning processes vary as a function of epistemic beliefs and contexts: Mixed method evidence from eye tracking and concurrent and retrospective reports. *Learning and Instruction*, *42*, 31–46.

Tuckman, B. W., & Kennedy, G. J. (2011). Teaching learning strategies to increase success of first-term college students. *Journal of Experimental Education*, *79*, 478–504.

U. S. Department of Education. (2016). *First in the world program*. Retrieved from https://www2.ed.gov/programs/fitw/index.html

Vandevelde, S., Van Keer, H., Schellings, G., & Van Hout-Wolters, B. (2015). Using think-aloud protocol analysis to gain in-depth insights into upper primary school children's self-regulated learning. *Learning and Individual Differences*, *43*, 11–30.

Veenman, M. V. J. (2007). The assessment and instruction of self-regulation in computer-based environments: A discussion. *Metacognition and Learning*, *2*(2–3), 177–183.

Veenman, M. V. J. (2011a). Alternative assessment of strategy use with self-report instruments: A discussion. *Metacognition and Learning*, *6*, 205–211.

Veenman, M. V. J. (2011b). Learning to self-monitor and self-regulate. In R. Mayer & P. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 197–218). New York, NY: Routledge.

Veenman, M. V. J., Elshout, J. J., & Groen, M. G. M. (1993). Thinking aloud: Does it affect regulatory processes in learning? *Tijdschrift voor Onderwijsresearch*, *18*, 322–330.

Weinstein, C. E., & Acee, T. W. (2013). Helping college students become more strategic and self-regulated learners. In H. Bembenutty, T. J. Cleary, & A. Kitsantas (Eds.), *Applications of self-regulated learning across diverse disciplines* (pp. 197–236). Charlotte, NC: Information Age Press.

Whipp, J. L., & Chiarelli, S. (2004). Self-regulation in a web-based course: A case study. *Educational Technology Research & Development*, *52*(4), 5–22.

Wingate, U. (2007). A framework for transition: Supporting "learning to learn" in higher education. *Higher Education Quarterly*, *61*, 391–405.

Winne, P. H. (2018). Cognition and metacognition within self-regulated learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 36–48). New York, NY: Routledge.

Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. Schunk & B. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Mahwah, NJ: Lawrence Erlbaum.

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.

Winters, F. I., & Azevedo, R. (2005). High-school students' regulation of learning during computer-based science inquiry. *Journal of Educational Computing Research*, *33*, 189–217.

Winters, F. I., Greene, J. A., & Costich, C. M. (2008). Self-regulation of learning within computer-based learning environments: A critical analysis. *Educational Psychology Review*, *20*, 429–444.

Young, D. G., & Hopp, J. M. (2014). *2012–2013 National survey of first-year seminars: Exploring high-impact practices in the first college year* (Research Report No. 4). Columbia: University of South Carolina, National Resource Center for the First-Year Experience and Students in Transition.

Zepeda, C. D., Richey, J. E., Ronevich, P., & Nokes-Malach, T. J. (2015). Direct instruction of metacognition benefits adolescent science learning, transfer, and motivation: An in vivo study. *Journal of Educational Psychology*, *107*, 954–970.

Zerr, R. J., & Bjerke, E. (2016). Using multiple sources of data to gauge outcome differences between academic-themed and transition-themed first-year seminars. *Journal of College Student Retention: Research, Theory & Practice*, *18*, 68–82.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, *48*, 135–147.

Manuscript received February 28, 2018
Final revision received June 17, 2019
Accepted June 27, 2019