Joel Sloan, UNITED STATES AIR FORCE ACADEMY, joel.sloan@afacademy.af.edu
Timothy Frank, AIR FORCE INSTALLATION AND MISSION SUPPORT CENTER, timothy.frank@us.af.mil
Lauren Scharff, UNITED STATES AIR FORCE ACADEMY, lauren.scharff@afacademy.af.edu
Karin Becker, UNITED STATES AIR FORCE ACADEMY, kbecker630@gmail.com

Knowledge Surveys: An Effective and Robust Student Self-Assessment and Learning Tool

ABSTRACT

Knowledge surveys (KS) are a student self-assessment tool consisting of an ability statement for each learning objective in a course. Students respond by rating their confidence in performing a specified skill. Pre-unit KS transparently communicate learning objectives, alert faculty to self-assessed knowledge and skills students possess as they enter a unit of instruction and support metacognitive learning through the unit's lessons and activities. Post-unit KS allow comparison of student self-assessed learning with faculty assessments of student learning, providing feedback to support continued development of self-assessment skills, as well as metacognitive and self-regulated learning skills. This study synthesizes several semesters of KS implementation across six different engineering courses, taught by six instructors utilizing multiple assessment types (exams, technical writing, design projects) and delivery modes (in person, hybrid, online). Results indicate that student KS scores aligned with faculty assessments, and alignment improved with cycles of performance and feedback in a subset of courses with similar assessment types. Student feedback indicates value in KS as a learning guide and as a helpful addition to these courses for learning and self-assessment, with student quotes demonstrating positive student learning practices and evidence of metacognitive thinking. Faculty find KS easily manageable; they also appreciate how KS help maximize alignment within their courses and provide additional data to support student learning. Knowledge surveys are thus an effective, robust, and relatively easy-to-implement method to systematically incorporate a self-assessment component into a variety of courses.

KEYWORDS

self-assessment, knowledge surveys, metacognition, self-regulated learning, spaced retrieval

INTRODUCTION

Self-assessment is a critically important part of learning (e.g., Andrade and Valtcheva 2009). As students navigate the learning process, they need to be self aware of when they "get it" and when they need more help. Performance on formative and summative assessments created by faculty provide feedback to students with respect to their achievement of outcomes, but ultimately, students need to develop the skill and habit of self-assessment in order to support academic success and life-long learning. Unfortunately, many students do not take the time to independently develop their self-assessment skills unless faculty provide a systematic structure to cultivate them. In this paper, we share a self-assessment tool, knowledge surveys (KS), that can serve as an effective, robust, and easily implemented method by which to engage students in self-assessment, building this important skill for learning.

Benefits of self-assessment to support learning

Supporting the development of self-assessment is somewhat akin to the age-old analogy of giving a person a fish versus teaching them how to fish. Deliberately developing student self-assessment skills aligns with the claim that faculty should be teaching more than just subject matter content; we should be teaching students how to learn (Sloan and Scharff 2022; Wirth and Perkins 2013).

Self-assessment has been linked to student self-regulation of learning and the development of metacognitive skills (Sullivan and Hall 1997; Wirth et al. 2021), as successful learners monitor and regulate their own learning in a formative process (Pintrich 2000). Formative assessment is important to deep learning (Rushton 2005), and theories of formative assessment often include a mandatory self-assessment component (e.g., Taras 2002). Andrade (2010) further asserts that the students themselves, via their own self-regulation and self-assessment, are the definitive source of feedback in formative assessment. Not only is self-assessment important in formal education, but it is also a key component of future professional development and lifelong learning (e.g., Bourke 2018).

Self-assessment requires reflection on one's perceived abilities and level of understanding. Metacognition involves monitoring these skills to develop self awareness, and then using that awareness to guide subsequent behaviors (Flavell 1976). The metacognitive skills practiced in student self-assessment include awareness focused on a process (in this case, the learning process), coupled with the intention to use that awareness to adjust as necessary if learning is deemed incomplete. When paired with instructor assessment and feedback, self-assessment can lead to an increase in self-efficacy (Tanner 2012) and achievement (Ross 2006) as the student understands how their actions, behaviors, and learnings contribute to performance.

Knowledge surveys as a metacognitive self-assessment tool

Although a vital part of learning, self-assessment often appears as an amorphous and intangible process without defined structure (Taras 2010). Nuhfer (1996) was an early advocate of knowledge surveys as a straightforward tool for student self-assessment. KS are collections of questions that support students' self-assessment of their understanding of course material and their skills and abilities across Bloom's levels of the cognitive domain (Clauss and Geedey 2010; Nuhfer and Knipp 2003). Bloom et al. (1956) identified a taxonomy of educational objectives with six progressive levels of expertise required for knowledge-based skills (remember, understand, apply, analyze, evaluate, and create). KS questions typically include explicit identification of the level of Bloom's domain at the end of each question, providing a clear communication of learning expectations to students. KS can be used to assess all levels of the taxonomy (Clauss and Geedey 2010; Nuhfer and Knipp 2003). We will explain more about development of KS questions in the methodology section of this paper; examples of KS questions and response options are also found in Beauregard, Sloan, and Brannon (2019); Nuhfer and Knipp (2003); Sloan and Scharff (2022); and Wirth and Perkins (2005).

Well-designed KS questions are specific and reflect the learning objectives within a course. They therefore serve as a tool to gauge alignment of learning objectives with assessments. Students typically complete the KS questions prior to each unit of the course (pre-instruction), and then also immediately prior to the summative assessment, such as exams, design projects, or writing assignments (post-instruction). When answering the questions, students rate themselves on their ability to answer the question, as opposed to providing a full response to the question. Instructors encourage students to use their self-assessment results to guide their subsequent learning behaviors consistent with Flavell (1976); if they are not confident in their ability to completely answer a question, they should revisit the material or seek other forms of help. Thus, KS support metacognitive learning because students focus their awareness on the extent of their learning and then use that awareness to guide their subsequent learning efforts (Nuhfer and Knipp 2003; Wirth and Perkins 2005; Wirth et al. 2021).

After an exam or other assessment, timely, constructive feedback comparing student self-assessments and performance on graded events can give students further opportunities to improve metacognitive skills (Zell and Krizan 2014). For example, when shown their post-KS scores alongside their actual performance on an exam or other assessment, students reflect about the differences and calibrate their own assessment abilities (Lindsey and Nagel 2013). This metacognitive calibration step can more powerfully guide subsequent learning than results from the instructor-graded assessment alone (Wirth et al. 2021). These KS implementation steps address several of the choke points and pitfalls in learning pointed out by Chew (2021), most notably the metacognitive pitfall of overconfidence, for which he proposes addressing via self-assessment and use of feedback.

KS as an effective tool for self-assessment

KS have been implemented in a variety of undergraduate and graduate courses across many disciplines (e.g., Nuhfer and Knipp 2003; Sloan and Scharff 2022). Most studies have focused on comparison of self-assessment data with grades on exams, although self-assessment mechanisms have been used for writing assignments (Becker and Sloan 2023; Sloan, Becker, and Frank 2022), undergraduate research experiences (McDevitt et al. 2016), and engineering design (Beauregard, Sloan, and Brannon 2019). These studies demonstrate that student self-assessments are aligned with faculty assessments (Clauss and Geedey 2010; Sloan and Scharff 2022; Wirth and Perkins 2005). Further, most first-year higher education students show good self-assessment accuracy when there is alignment between the KS and the actual assessments (Nuhfer et al. 2016a). Overall, KS are not about an end goal of student conformity to the instructor grade but about giving students agency and control over their own learning through metacognition and self-regulation. Thus, it seems reasonable to incorporate KS as a pedagogical tool in order to encourage student engagement in and development of self-assessment skills.

Learning benefits of KS for students

Prior research demonstrates multiple benefits of integrating KS for student learning. Students complete pre-unit KS as a cognitive heads up on the material to follow in the course and can use the KS as a learning guide as they navigate each block of a course (Sloan and Scharff 2022). The KS questions clarify instructors' learning expectations and forecast the challenge ahead (Wirth et al. 2021), which supports the development of self-regulated learning skills (Mahlberg 2015). The use of both pre-unit and post-unit KS support spaced retrieval practice and enhances long-term learning by encouraging students to reflect on their ability to answer the questions multiple times (Ariel and Karpicke 2018; Chew 2021). Finally, KS have the potential to support other behaviors that are believed to support learning, such as focusing attention (Chew 2021), driving learners toward taking greater responsibility for their own learning (Falchikov and Boud 1989), challenging students to think more deeply about a subject (Sullivan and Hall 1997), and developing students into life-long learners (Fink 2003).

Teaching benefits for faculty

In addition to supporting students' development of their self-assessment skills, KS provide a variety of benefits for faculty (e.g., Wirth and Perkins 2005). Prior to the start of the term, the incorporation of KS within a course can heighten faculty awareness of alignment between lesson goals, activities, and assessments. As the course unfolds, results from student pre-unit KS provide valuable data, as they allow the faculty member to ascertain the level of understanding of a particular cohort of students on a topical basis and use that to guide instruction. Faculty who are attuned to their students' feedback can learn from the student perspective and use class time more efficiently. For example, if students already have a moderate or high degree of self-assessed understanding on a particular topic, then the faculty member can spend less time on that topic than if students had low self-assessed understanding (Nuhfer and Knipp 2003; Wirth and Perkins 2005). Alternatively, individual students self-assessing at a lower level in the pre-survey may need additional instruction and activities to catch up with the rest of the class. In this way, KS can support metacognitive instruction practices of faculty. On an even more targeted level, KS data can support individualized conversations with students about their learning strategies and performance (Sloan and Scharff 2022). Finally, KS can also serve as a means to evaluate the effectiveness of new pedagogical approaches (Wirth and Perkins 2005).

CONTEXT AND PURPOSE

Given the potential benefits and efficacy of KS, several faculty members at the United States Air Force Academy (USAFA) initiated an interdisciplinary scholarship of teaching and learning (SoTL) research project in the spring 2018 semester focused on student self-assessment using knowledge surveys. Approved through the institutional review board, the overall project included dozens of faculty members participating in courses that span the humanities, social sciences, basic sciences, and engineering. The goals of the overall project were to answer three research questions: 1) How aligned are student self-assessments with faculty assessments? 2) What are the student attitudes and self-reported behaviors related to KS? and 3) What are the instructor attitudes, decisions, and lessons learned regarding the implementation of KS in their courses?

In the process of answering these research questions in other works (e.g., Sloan and Scharff 2022; Scharff et al. in press), we also discovered that KS are robust, effective, and relatively easy to implement. The purpose of this paper is to share a synthesis of our results implementing KS across a variety of civil engineering courses to support our claims that KS offer a robust and manageable method by which to support best practices for teaching and learning, including student self-assessment and metacognitive skills. Derived from our extensive implementation of KS in engineering courses, the key contributions of this paper are the synthesis of results across many different courses from second-year to fourth-year students in a single program of study, by faculty with differing education and experience levels, using several different assessment types (exams, design projects, and technical writing), while providing longitudinal results in one course that includes in-person, online, and hybrid delivery modes.

METHODOLOGY

This study builds on the work by Beauregard, Sloan, and Brannon (2019); Sloan and Frank (2023); Sloan and Scharff (2022) at USAFA (Sloan, Beauregard, and Russell 2023). In total, our study includes KS data from 18 sections of six engineering courses from 2018 to 2022. Six faculty taught 219 total students across these courses. The following subsections describe the institutional and program context, the courses and instructors, and the implementation methods.

Institutional context

USAFA is an undergraduate-only military-focused institution that requires students to complete their course of study within four years. Students complete a large general education curriculum, totaling 93 semester hours across the humanities, social sciences, basic sciences, and engineering. The demographics of the students in the courses in this study were 82% male and 18% female, with 67% White, 11% Hispanic, 9% Asian, 7% Black, 1% American Indian, 1% Native Hawaiian/Pacific Islander, and 4% not specified.

Faculty

As an undergraduate-only institution, faculty primarily focus on teaching. Class sizes are generally small, with civil engineering major's courses typically having 10–18 students in each section. All faculty in the study voluntarily used KS in their courses: three were male, three were female, and all six were White. Three held a PhD as their highest degree, and three held a master of science (MS) degree. Teaching experience of the faculty varied widely, with three being categorized as "new" (zero to two years' experience), two being characterized as "mid" (three to 10 years' experience), and one being categorized as "advanced" (11+ years of experience). In some cases, faculty used KS over multiple courses and multiple years, so experience grew during the time of the study.

Courses

Courses in this study included one course from the general education curriculum, two third-year courses required for students' major, a fourth-year required course, a major's elective course, and a design elective course. Table 1 summarizes course titles, the nature of the course (general education, elective major's course, or required major's course), the assessments, and the number of KS questions in each course.

Faculty implemented self-assessment using KS alongside a variety of course assessment types. In most of our courses, the students' KS questions focused on the content within each block of instruction, and faculty directly compared KS scores with instructor grades on exams. As engineering courses, most of the exams consisted of problems at the "apply" to "create" levels of Bloom's taxonomy, where students used principles from the courses to solve problems. However, within computer applications, the KS questions focused on aspects of technical writing. Students repeated the same 11 KS questions while completing the technical writing assignment for each of the four blocks of instruction. Also unique from the other courses, in foundation engineering, assessment 2 KS questions focused on their design project, offering comparisons with instructor grades on that project. Thus, our results compare instructor assessments of student performance with student self-assessments across assessment types that include exams, technical writing, and design projects, as shown in Table 1.

Table 2 summarizes the six courses, semester offered, instructor education and experience, number of students, course, and course delivery mode. Data come from a total of 219 students, resulting in 720 data pairs comparing KS and faculty assessed student performance. Importantly, faculty shared grading of assessments among those teaching the same course in the same semester, which helps to reduce the opportunity for grading bias from a single instructor.

Table 1. Summary of course and assessment characteristics

Course	Nature of course	Primary student population	Assess. 1 (# of KS questions)	Assess. 2 (# of KS questions)	Assess. 3 (# of KS questions)	Assess. 4 (# of KS questions)
Fundamentals of Mechanics	General education	Second-year	Exam (18)	Exam (13)	Exam (26)	-
Computer Applications	Elective CE major's course	Third-year	Technical writing (11)	Technical writing (11)	Technical writing (11)	Technical writing (11)
Fundamental Hydraulics	Required CE major's course	Third-year	Exam (22)	Exam (23)	Exam (19)	Cumulative final exam (89)
Analysis of Structures	Required CE major's course	Third-year	Exam (11)	Exam (9)	Exam (9)	-
Foundation Engineering	Elective CE major's course	Fourth-year	Exam (42)	Design project (25)	-	-
Pavement Design	Required CE major's	Fourth-year	Exam (88)	Exam (73)	Cumulative final exam (221)	-

Table 2. Course offerings chronologically by semester with instructor(s), enrollment, and delivery mode

Course	Semester(s)	Instructor(s) education and experience*	Number of students	Course delivery mode
Pavement Design	Spring 2018	PhD - mid; MS - new	35	In-person
Foundation Engineering	Fall 2018	PhD - mid	6	In-person
Fundamental Hydraulics	Spring 2019	PhD - mid	16	In-person
Fundamental Hydraulics	Fall 2019	PhD - mid; MS - new	32	In-Person
Fundamental Hydraulics	Spring 2020	PhD - mid	16	In-person / online
Fundamental Hydraulics	Fall 2020	PhD - mid	23	Hybrid
Fundamental Hydraulics	Spring 2021	PhD - mid	30	Hybrid
Computer Applications	Fall 2021	PhD – mid; PhD - advanced	21	In-person
Analysis of Structures	Fall 2021	PhD - mid; MS - new	21	In-person
Fundamentals of Mechanics	Fall 2021	PhD - mid	13	In-person
Analysis of Structures	Spring 2022	PhD - mid	6	In-person

*Experience legend: New (zero to two years teaching); mid (three to10 years teaching); advanced (more than 10 years teaching)

Faculty taught all courses from Table 2 in a single delivery mode, except for Fundamental Hydraulics. Given that the offerings of Fundamental Hydraulics spanned the COVID-19 pandemic, faculty taught in a variety of delivery modes. Faculty taught pre-pandemic sections (spring and fall

2019) fully in-person. When the pandemic hit in the spring 2020 semester, faculty taught approximately 60 percent of the course in-person before shifting to synchronous online instruction for the remaining 40 percent of the course. Faculty delivered the fall 2020 and spring 2021 sections in a hybrid manner. Hybrid delivery consisted of asynchronous online content lessons with each 53-minute lesson condensed into a 15–20 minute video that students watched before working example problems reinforcing lesson content. Every three to four lessons, the instructor met in-person with the students, and these lessons consisted entirely of board-work (problem solving at the "apply" Bloom's level) that enabled the instructor to monitor students' grasp of the concepts in a formative way prior to the exams.

KS implementation procedures

For most courses, faculty developed KS questions from the existing learning objectives by simply adding an "I can" ability statement to the beginning of the objective. Having pre-existing learning objectives makes the transition to KS relatively easy, presuming objectives are well-written and specific (not general).

Figure 1 illustrates sample KS questions from three of the courses in this study with Bloom's level shown in parenthesis at the end of each question. Bloom's level varied within courses, with upper-level courses having more questions at higher Bloom's levels.

We also examined Cronbach's (1951) alpha for the KS in each course (29 to 221 questions per course other than Computer Applications which had 11 questions as shown in Table 1) to ensure they had good internal consistency at the pre- and the post-unit implementations. Overall, the internal consistency check assured us that we had a reliable self-assessment tool as pre-unit KS Cronbach's alpha ranged from 0.79 to 0.95 and post-unit Cronbach's alpha ranged from 0.79 to 0.98. These results align with those from Nuhfer and Knipp (2006), who report that reliability for KS are generally high.

KS implementation procedures across all courses and semesters from Table 3 generally followed processes described by Sloan and Scharff (2022). Researchers designed these common procedures to reinforce the value of KS to students (practicing an important skill, not simply busywork). They provide the feedback necessary for students to improve their self-assessment skills with each successive use. The procedures are summarized as follows:

- Faculty introduced KS and Bloom's taxonomy during lesson 1. Course policy letters contained information describing KS and Bloom's taxonomy and explained why self-assessment is an important skill to develop.
- Students had continuous access to KS questions to use as a learning guide in each course, often via multiple means: printed KS questions from lesson 1, a document listing the KS questions on the course learning management system (LMS), and an electronic survey on the course LMS that allowed unlimited attempts throughout the semester.
- Students completed either:
    - A comprehensive pre-course KS at the beginning of the course (2018 and 2019 offerings) or
    - Pre-unit KS at the beginning of each unit. We opted for the pre-unit KS method in later offerings as it supports spaced retrieval practice in closer succession with the learning activities and is likely more effective (Sloan and Scharff 2022).
- KS questions were specific and aligned with explicit content and instruction (Pintrich 2002). The KS questions corresponded one-for-one with the learning objectives for the lessons that made up a given block of instruction.

Figure 1. Sample KS questions and response options from three courses in this study

Fundamentals of Mechanics:

- I can resolve a force vector into orthogonal components (Apply).
- I can solve for support reactions using a free-body diagram and the summation of forces and moments (Apply).
- I can calculate the maximum tensile and compressive normal stress due to bending (Apply).

Response options:

a) I am unable to perform the task.
b) I can begin to perform the task but am quickly overwhelmed.
c) I can make progress toward performing the task but fall well short of completing it.
d) I can almost completely perform the task.
e) I am able to perform the task for the exam.

***Computer Applications* (Technical Writing):**

- I am able to fully consider the audience and their needs as I write (Evaluate).
- I am able to communicate ideas in writing using simple, direct sentences (Apply).
- I am able to integrate visuals (e.g. charts, maps, graphs, tables, images) with informative labels into a technical document (Apply).

Response options:

a) I have no proficiency in this task at this time.
b) I have very limited proficiency in this task at this time.
c) I have some proficiency in this task at this time.
d) I have good proficiency in this task at this time.
e) I have excellent proficiency in this task at this time.

Pavement Design:

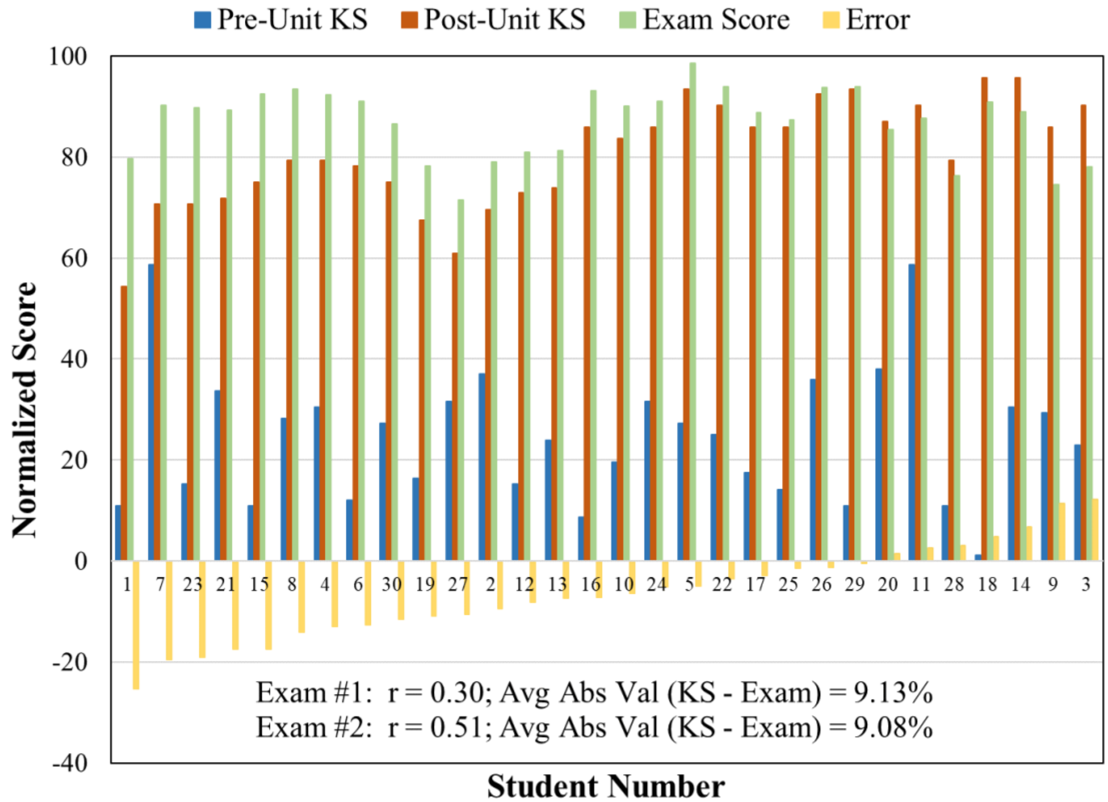
- I can explain the role of and the location of reinforcement in concrete pavements (Understand).
- Given traffic information, I can calculate the total number of equivalent single axle loads (Apply).
- I can design a flexible pavement using the 1993 AASHTO asphalt design procedure and the structural number concept (Create).

Response options:

a) I am unable to perform the task at this time.
b) I am able to partially perform the task at this time.
c) I can perform the task for evaluation at this time.

- Students also completed a post-unit KS for each assessment (exam, design project, or writing assignment, as noted in Table 1) within the 24 hours prior to completing the assessment. Most instructors offered a small number of points on each assessment event (typically 5%) to 1) communicate to students the value of completing the KS, 2) reward them for their time to complete it, and 3) as an incentive for students to complete the KS. KS completion rates in each course exceeded 95%.
- Although not required, faculty encouraged students to consult the KS regularly throughout the course as a learning guide rather than solely in preparation for assessments.
- Faculty provided feedback to students after each assessment on how their self-assessment compared with the faculty assessments. An example of such feedback from Fundamental Hydraulics is given in Figure 2, which depicts pre-unit KS, post-unit KS, assessment score, and error (KS minus exam score) for each student. Students knew their own student number but could see anonymous data from other students. Faculty challenged the students to self-assess accurately and described the goal of self-assessment error being within +/-10% (Nuhfer et al. 2016b).
- Instructors also shared the correlation between KS and assessment scores and the average absolute value of error of the cohort for that exam. As the semester progressed, these correlation and error statistics were shared for the previous as well as the current exam, to communicate whether the class was improving in their self-assessment accuracy. Longitudinal data from individual students also enabled conversations between faculty and students about why a student, for example, may be habitually over- or under-confident.

Figure 2. Sample feedback given to students after exam 2 in Fundamental Hydraulics (spring 2021)



This figure shows pre-unit KS, post-unit KS, exam score, and error (KS minus exam score) by student, and the correlation of KS to exam score and average absolute value of KS minus exam score for the student cohort.

Student feedback surveys

To better understand the student perspectives, faculty collected student feedback at mid and end-of-semester by one of two methods 1) via a paper survey administered by a neutral third party, or 2) via an online survey. Faculty teaching the respective courses did not have access to the student survey results until after they submitted final grades for the semester. The surveys included the following closed-ended questions to which students responded using a 4-point Likert scale (strongly disagree, disagree, agree, strongly agree):

1) The knowledge surveys provide clear guidance as to what I am expected to know or be able to do for the exam or design project.
2) The knowledge surveys are a useful addition to this course with respect to supporting my learning of the course material.
3) The knowledge surveys are a useful addition to this course with respect to supporting my ability to self-assess my knowledge and preparation for exams.

In addition, the feedback survey included two open-ended questions:

1) Explain why the knowledge surveys are or are not useful. Please include any recommendations for improvement.
2) How would you suggest future students use the knowledge surveys to succeed in this class?

RESULTS

Our evidence of effectiveness, robustness, and ease of implementation includes data from two sources that span the variety of courses and instructors: 1) the alignment of student self-assessments with faculty grades (i.e., amount of error) and 2) student perspectives on the use of KS and their value as a self-assessment tool.

Assessing KS effectiveness and student development of self-assessment skills

Histograms of error

Using a format inspired by Nuhfer et al. (2017), Figure 3 shows histograms summarizing error (KS self-assessment minus actual assessment score) for all assessments within a given course. Error closer to zero indicates better alignment of student self-assessments with faculty grades. A positive number indicates students are over-confident (KS higher than assessment score) compared to their grade, and a negative number indicates students are under-confident. Tighter grouping of error around zero indicates that student self-assessments are better aligned with faculty grades.

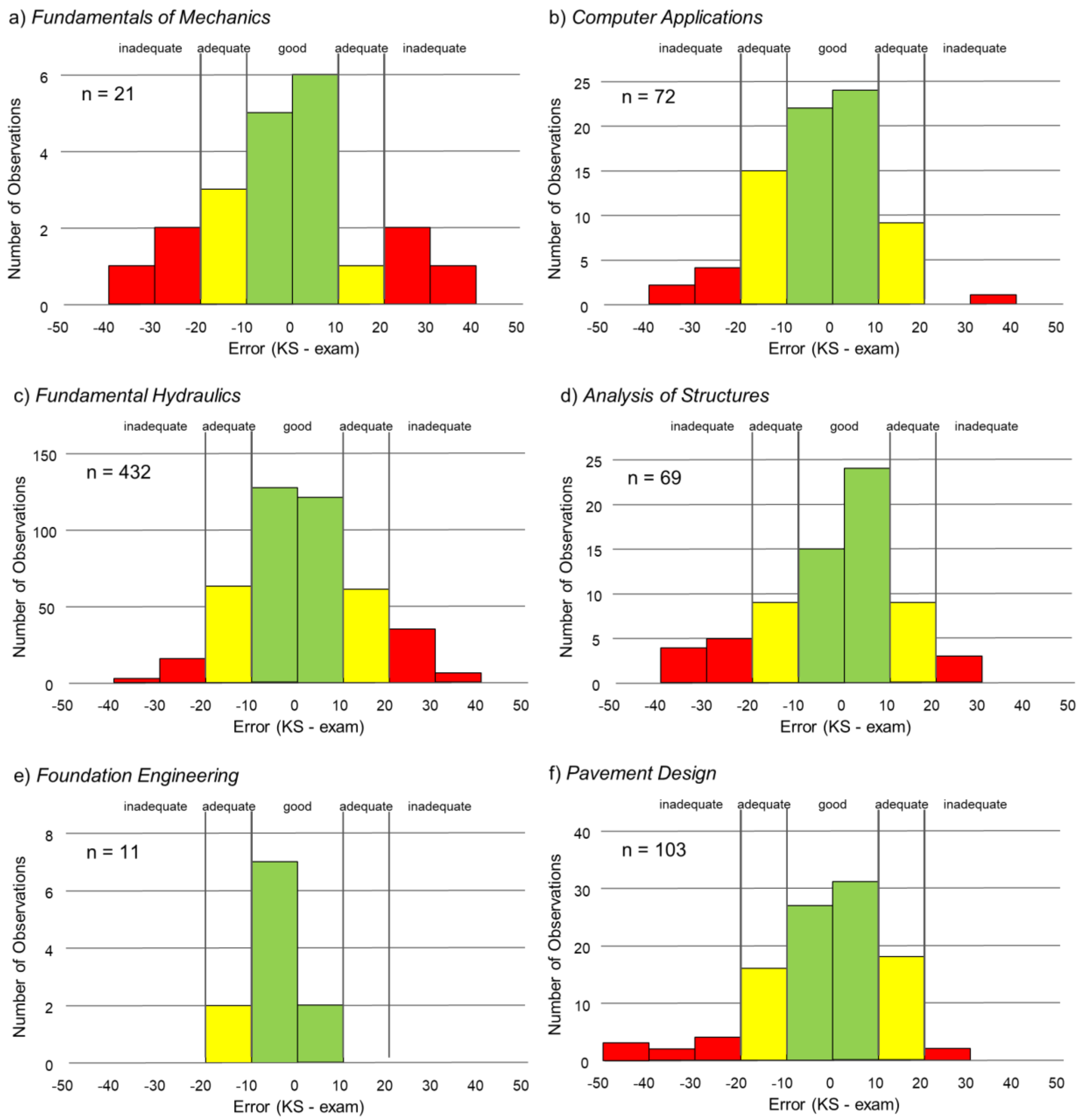
The use of the term "error" rather than "self-assessment error" is intentional, as error may be due to a variety of reasons. They could be due to a student's lack of self-assessment ability but could also result from a lack of clear learning objectives, misaligned courses or assessments, bias in grading, or unclear expectations on the part of the faculty member. Thus, excessive error, especially if shown across multiple students, should be a clue to faculty to delve deeper into their own possible contributions to the error.

We chose to label a range of +/-10% as "good," +/-10 to 20% as "adequate," and greater than +/-20% as "inadequate." Although there is no standard convention, our selection of these ranges is consistent with (but a simplified version of) Nuhfer et al. (2017). Note that the histograms include multiple responses from students enrolled in each course, because each course had multiple assessments (see Table 1). Later in this sub-section, we address potential improvement in self-assessment skills across assessments.

For all courses, some of which multiple faculty members taught, the histograms in Figure 3 reveal a distribution of error with a mean near zero, indicating that student self-assessments using KS are generally well-aligned with instructor assessments. Although not presented in this paper, we did review results for individual instructors with the results for all instructors showing a similar theme with alignment of error centered around zero. Across the six courses, more than half (56%) of KS scores were of good accuracy (within +/- 10% of the faculty grade) while only 16% were inadequate (outside of +/- 20% of the faculty grade). Overall, these results indicate that most students in these six courses effectively practiced and exercised their self-assessment skills using KS.

Given that Fundamental Hydraulics was the only course in this study where faculty taught in more than one learning mode, we used it to take a closer look at the data across the different learning modes: in-person during spring and fall 2019, a combination of in-person and online during spring 2020, and hybrid during fall 2020 and spring 2021. Figure 4 shows histograms of error for these three learning modes. Students in each of the three modalities showed similar patterns of self-assessment, with most students self-assessing within the good and adequate categories, indicating that student self-assessments via KS align with faculty assessments, regardless of learning mode.

Figure 3. Histograms of error (student KS minus faculty grade) on each assignment by course



Correlations between student KS and faculty assessment

The histograms above aggregated data across assessments within each course in Figure 3; however, correlations between student KS and faculty assessments can also help us examine the alignment of self-assessments and grades for individual courses and sections of courses. Earlier we claimed that KS can serve as a metacognitive tool to support students in the development of their self-assessment skills. To examine this claim, we narrowed the data set to courses with multiple, similar assessment types within the semesters. Thus, for this analysis, we focus on results from three courses (Fundamental Hydraulics, Analysis of Structures, and Fundamentals of Mechanics), each of which had three exams where we could assess students' self-assessment skills across similar assessments.

Figure 4. Histograms of error by learning mode in Fundamental Hydraulics

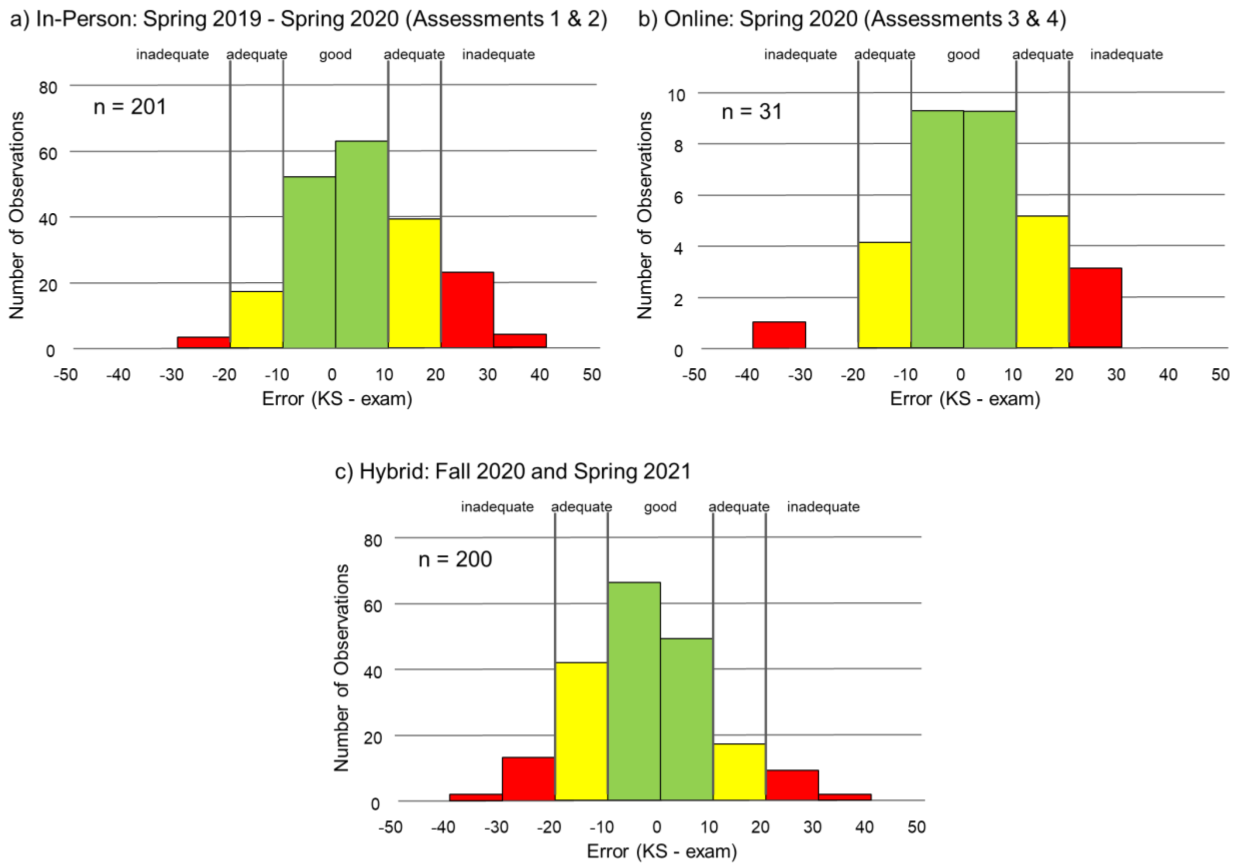


Table 3 contains correlations for a combination of the three courses taught by four of the six faculty in this study over eight semesters. In this case, the course was not the variable of interest but rather the exam sequence. Thus, we combined the data across courses for exams 1, 2, and 3 and calculated a combined correlation for each exam. All correlations are positive and statistically significant ($p < 0.001$), and there is a clear increase in the correlation between student KS and instructor grades from exam 1 to 2 to 3 in these three courses. These results illustrate that alignment of student self-assessments with faculty assessment of student performance can improve with practice and feedback through a given semester, and they provide evidence that students are applying the feedback they receive on their self-assessment ability (similar to Figure 2) in a metacognitive way to influence subsequent behaviors, consistent with Flavell's (1976) definition and results from the literature (Sloan and Frank 2024).

Student feedback data

Student responses to Likert-scale feedback questions

The 4-point Likert-scale questions provide insight into student perceptions about the utility of KS. Average responses and standard error for each course on a 4-point Likert scale are shown in Table 4. Data from Pavement Engineering in the spring 2018 semester is unavailable as we developed and implemented these questions in the fall 2018 semester. Student feedback is generally positive, irrespective of the course or an individual instructor. The average response is greater than 3.0 (agree) across all courses (N = 112) and is above 3.0 for most questions in most courses.

Table 3. Correlations between KS and instructor assessment in courses with three consecutive unit exams

Courses	Correlation between KS and instructor exam 1 (df)	Correlation between KS and instructor exam 2 (df)	Correlation between KS and instructor exam 3 (df)
Fundamental Hydraulics, Analysis of Structures, and Fundamentals of Mechanics	0.29** (135)	0.32** (140)	0.44** (135)

df = degrees of freedom (n – 2)
**statistically significant correlation ($p < 0.001$)

Table 4. Averages and (standard errors) of student feedback data using a 4-point Likert scale

Question	Fund. of Mech. (N = 3)	Computer Apps (N = 14)	Fund. Hydraulics (N = 80)	Analysis of Structures (N = 9)	Found. Engr. (N = 6)	Average response (N = 112)
Good learning guide	3.67 (0.272)	3.14 (0.094)	3.17 (0.077)	3.00 (0.105)	3.83 (0.152)	3.19 (0.058)
Useful addition for learning	2.67 (0.544)	3.14 (0.171)	3.11 (0.068)	2.86 (0.105)	3.33 (0.304)	3.09 (0.057)
Useful addition for self-assessment	2.33 (0.272)	3.07 (0.122)	3.14 (0.070)	3.00 (0.189)	3.83 (0.152)	3.13 (0.057)

Student responses to open-ended feedback questions

To analyze the open-ended student feedback questions, we conducted a thematic analysis of student responses. Two reviewers (not among the six faculty teaching the courses) independently read the student responses and developed categories of responses without any preconceived notions on what those categories might be. The reviewers reached an agreement on the categories. They then independently coded all the responses into these categories. The coders checked interrater reliability; if any category showed more than 10% disagreement, the coders would recalibrate and independently recode until all categories showed less than 10% disagreement (most ended up with less than 5% disagreement). In some cases, a single written response from a student had portions that touched on more than one category.

Response themes from the open-ended questions are summarized in the tree plots shown in Figures 5 and 6 which summarize 94 and 98 student responses (respectively) to open-ended questions. Responses are grouped by five major themes and in two categories: procedural (green) and metacognitive (blue), along with a representative quote from each theme. Percentages in each figure total more than 100%, as some student responses fell into more than one theme. Since feedback was voluntary, fewer students responded compared to the number of students in the courses from Table 2.

Responses to the question about KS usefulness, shown in Figure 5, were overwhelmingly positive. Students noted several procedural (green boxes) and metacognitive (blue boxes) benefits of KS, with the largest group of responses noting how KS helped them study. However, almost half of the students clearly noted metacognitive benefits, with quotes indicating that they used self-regulation to guide behaviors per Flavell's (1976) definition of metacognition.

Based on their experiences with KS in these courses, students also had positive suggestions for future students on how they should use KS, as shown in Figure 6. Like in Figure 5, responses in Figure 6 are also categorized by procedural (green boxes) and metacognitive themes (blue boxes), with more than half of students highlighting positive metacognitive benefits.

One student quote especially captured several of the procedural and metacognitive themes above:

> I actually enjoyed this as an aspect of the class as it added almost like a mini study guide for me to work through and remember what the particular survey covered. This allowed me to formulate a study plan to ensure any questions I felt deficient on were covered again adequately. I treated these as my form of study guide and I did not feel it necessary to work out many problem types over again when I understood them to my desired extent. I felt that these surveys closely followed the learning objectives too and asked something about almost every topic type so I felt prepared for a test when I used these actively.

In summary, the open-ended feedback responses from students reveal positive perspectives of KS and illustrate the metacognitive thinking and self-regulated learning skills that KS are designed to stimulate in students.

Figure 5. Open-ended question 1: Explain why the knowledge surveys are or are not useful

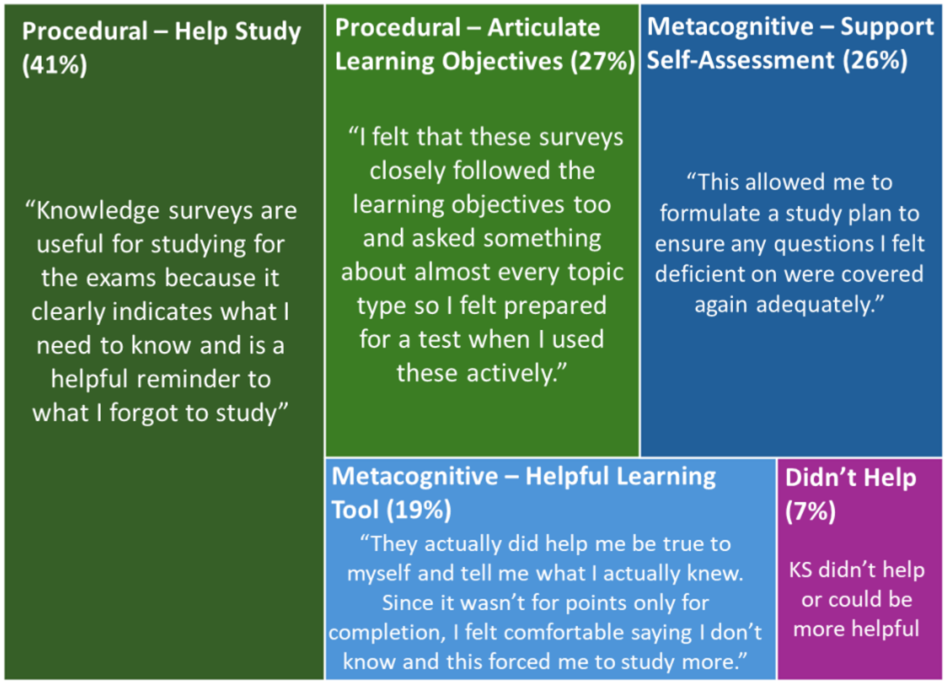
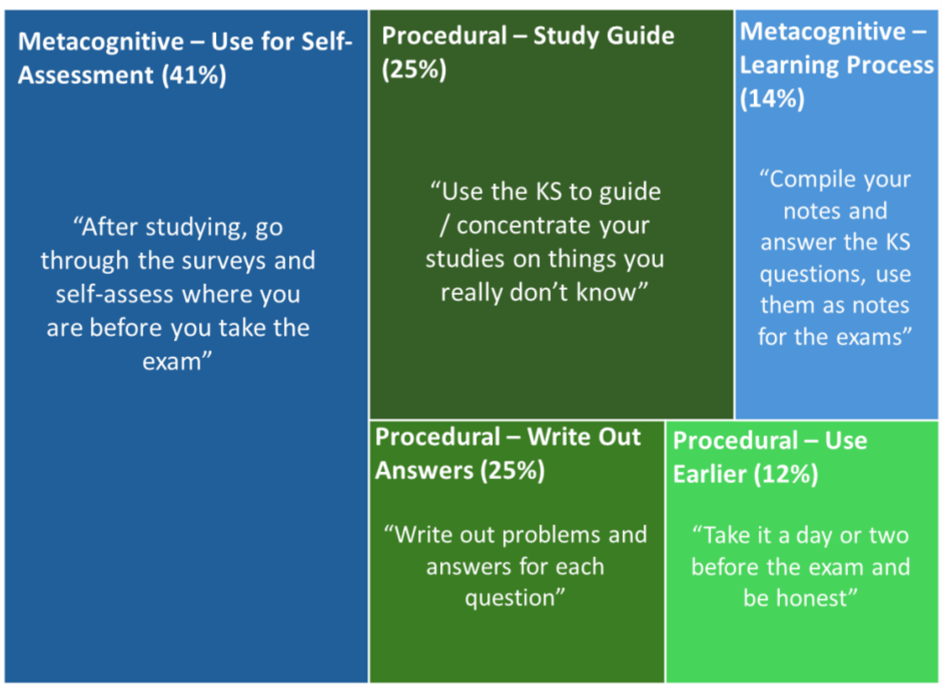


Figure 6. Open-ended question 2: How should future students use KS?



DISCUSSION

Overall, our results support our claim that KS are an effective and robust method for developing student self-assessment skills. Although prior work has shown elements of effectiveness (e.g., Nuhfer and Knipp 2003), the uniqueness of our work comes via demonstrating these qualities of KS in a breadth of courses, assessment types, delivery modes, and varying faculty education and experience levels. The histograms with error grouped near zero and the significantly positive correlations illustrate that KS are a reliable tool for engaging students in self-assessment in a way that aligns with instructors' articulation of learning objectives. Further, student feedback indicates that students generally look favorably on KS as a good learning guide and a useful addition to support learning and their self-assessment abilities. Their open-ended comments suggest that students exercised metacognitive skills when engaging with KS. These responses illustrate Flavell's (1976) conceptualization of metacognition, because our students used the KS to assess their confidence in a topic area and then take subsequent action to further study that topic in preparation for an assessment or to stop studying when they reached a level of competency on that topic. Further, the improved alignment between student KS and faculty assessments from exam 1 to exam 3 (Table 3) indicates that students can incorporate metacognitive feedback on their self-assessment skills and further develop these skills in the context of an academic course.

The robustness of KS stems from their effective implementation across courses, faculty members, assessment types, and learning modes. Essential elements of effective implementation in any context include: 1) specific KS questions, 2) alignment of learning objectives (from which KS questions are derived), learning activities, and assessments, and 3) clear feedback to students on how their KS compare with instructor assessments. Faculty in this study possessed varying levels of teaching experience and education and successfully implemented KS in six engineering courses. These courses encompassed a range of levels, topics, and delivery modes, and the assessments included exams, projects, and writing assignments. Finally, student perspectives also support the robustness of KS due to the positive perceptions of KS across all courses. Our results build on the findings of other researchers who have implemented KS across other disciplines (Clauss and Geedy 2010; Nuhfer and Knipp 2006; Wirth and Perkins 2005).

Faculty perspectives

As part of our work over the years, we and the other faculty involved at USAFA discussed our experiences and believe it is important to explore and share faculty perspectives. It takes time and effort to make changes in our courses, and it is important to consider faculty perspectives when promoting a pedagogical approach. We found that faculty across all levels of experience were able to implement KS without unreasonable levels of effort, plus they reported several areas of benefit.

Ease of implementation

As described in the methodology section, KS are developed by taking existing learning objectives from courses and rephrasing them as "I can" ability statements, as shown in Figure 1. Post-processing of KS data to produce a normalized self-assessment score and a plot-like Figure 2 does require additional time and effort for the first iteration (<2 hours in our experience), but it then becomes much quicker and easier to simply repeat these steps with each successive iteration of an assessment (<30 minutes). Faculty in this study were eager to do this postprocessing in order to determine whether student KS scores aligned with their grades and whether students improved in their self-assessment skills. More information regarding the implementation of KS is found in Nuhfer and Knipp (2003), Sloan and Scharff (2022), and Scharff et al. (in press).

KS support faculty efforts in course design and alignment

The implementation of KS forces faculty to continually evaluate the design of their courses and the alignment between learning objectives, in-class and out-of-class activities, and formative and summative assessments. Although most faculty have a broad understanding that course alignment is important, the deliberate act of developing the KS questions from learning objectives, designing learning activities to achieve the objectives, and evaluating student performance in light of their KS scores force faculty to examine this alignment more closely.

KS facilitate productive conversations with students and scaffold effective learning

Rather than simply having an assessment grade to discuss, KS also provide faculty with data from students' self-assessments. This helps lead productive and individualized conversations, because there are many possible combinations of KS and instructor assessment scores. For example, a low-performing but accurately self-assessing student may require feedback on motivation or on effective study and preparation techniques to improve their skills. An underconfident student may become more confident by viewing their high performance on an exam in relation to their self-assessment score (as in Figure 2) and by having a conversation with their instructor about self-assessment and academic performance. And an overconfident student may learn to calibrate future self-assessments by viewing feedback on their self-assessment scores in relation to instructor assessments as recommended by Chew (2021). Incorporating KS in a course provides a natural path for faculty to discuss learning strategies and the importance of self-assessment in their classes.

Further, most faculty appreciate that they are teaching students more than subject-matter content; they are also teaching students about the learning process itself and how to acquire new knowledge (Wirth and Perkins 2013). KS support spaced retrieval practice (Ariel and Karpicke 2018) starting when students complete the pre-unit KS at the beginning of each unit as a cognitive "heads up" on the learning material upcoming in that unit. Students view the KS questions (learning objectives) for each lesson and in a well-designed course, formatively self-assess using the KS questions through in-class activities and low-stakes assessments. Finally, students self-assess using the post-unit KS prior to the graded assessment. Self-assessing with the KS early and in continuous preparation for a summative assessment allows students to apply the metacognitive exercise of spaced retrieval practice.

Limitations

We obtained the results from this study at a single undergraduate-only institution in the United States with 1) a teaching focus and small class sizes (10—18 students in major's courses), 2) a strong culture of faculty development, and 3) a military focus. Although the relatively small sample sizes in each course precluded our ability to investigate effects of demographics, and the results are primarily limited to the civil engineering program of study, our sense, given the experience of other faculty at our institution and our review of the literature, is that the results of this study are applicable to other contexts and disciplines. A logical next step would be to analyze the potential impact of gender, race, and ethnicity on the accuracy of student self-assessment as our sample size continues to grow.

CONCLUSION

This paper shares results from incorporating knowledge surveys into six engineering courses taught by six faculty members with widely varying experience levels over nine semesters to 219 students in order to develop and evaluate student self-assessment skills. KS provided the means to encourage students to engage in self-assessment and cultivate metacognitive and self-regulated learning skills that they might not develop otherwise. Our results indicate that KS are robust, effective, and relatively easy to implement. KS can offer a positive addition to a course by providing a systematic way for students to cultivate self-assessment and metacognitive skills.

AUTHOR BIOGRAPHIES

Joel Sloan (USA) is the head of the Department of Civil and Environmental Engineering at the US Air Force Academy. He earned his PhD in civil engineering from Virginia Tech.

Timothy Frank (USA) serves at the Air Force Installation and Mission Support Center and is a former faculty member at the US Air Force Academy. He earned his PhD in civil engineering from Stanford University.

Lauren Scharff (USA) is the director for the Scholarship of Teaching and Learning at the US Air Force Academy. She earned her PhD in human experimental psychology from the University of Texas at Austin.

Karin Becker (USA) is the former director of the Communication Strategies Center at the US Air Force Academy. She earned her PhD in health communication and public discourse from the University of North Dakota.

ACKNOWLEDGMENTS

The authors would like to acknowledge Lauren Sahagun, Dirk MacDonald, and Paige Blackburn, who taught sections of courses in this study, as well as Will Atkins and Holli Meekins who helped to analyze the student feedback data in Figures 5 and 6.

DISCLAIMER

The views expressed in this paper are those of the authors and do not necessarily reflect the official policy or position of the United States Air Force Academy, the US Air Force, the US Department of Defense, or the US Government.

REFERENCES

Andrade, Heidi L. 2010. “Students as the Definitive Source of Formative Assessment: Academic Self-Assessment and the Self-Regulation of Learning.” *NERA Conference Proceedings* 2010: 25. https://opencommons.uconn.edu/nera_2010/25/.

Andrade, Heidi, and Anna Valtcheva. 2009. “Promoting Learning and Achievement through Self-Assessment.” *Theory Into Practice* 48 (1): 12–19. https://doi.org/10.1080/00405840802577544.

Ariel, Robert, and Jeffrey D. Karpicke. 2018. “Improving Self-Regulated Learning with a Retrieval Practice Intervention.” *Journal of Experimental Psychology Applied* 24 (1): 43–56. https://doi.org/10.1037/xap0000133.

Beauregard, Melissa S., Joel A. Sloan, and Michael Brannon. 2019. “Summarizing the Use of Knowledge Surveys to Inform Effective Learning and Teaching Practices.” *Proceedings of Rocky Mountain Section of the American Society for Engineering Education’s Annual Conference*, May 2019, Laramie, WY.

Becker, Karin L., and Joel A. Sloan. 2023. “Implementing Technical Communication Instruction in a Civil Engineering Course: The Value of Knowledge Surveys and Peer Review.” *Journal of Civil Engineering Education* 149 (4). https://doi.org/10.1061/JCEECD.EIENG-1927.

Bloom, Benjamin S., Max D. Englehart, Edward J. Furst, Walker H. Hill, and David R. Krathwohl. 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain.* New York: David McKay Company.

Bourke, Roseanna. 2018. “Self-Assessment to Incite Learning in Higher Education: Developing Ontological Awareness.” *Assessment & Evaluation in Higher Education* 43 (5): 827–39. https://doi.org/10.1080/02602938.2017.1411881.

Chew, Stephen. 2021. "An Advance Organizer for Student Learning: Choke Points and Pitfalls in Studying." *Canadian Psychology* 62 (4): 420–27. https://doi.org/10.1037/cap0000290.

Clauss, Jon, and Kevin Geedey. 2010. "Knowledge Surveys: Students Ability to Self-Assess." *Journal of the Scholarship of Teaching and Learning* 10 (2): 14–24.

Cronbach, Lee J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16 (3): 297–334. https://doi.org/10.1007/BF02310555.

Falchikov, Nancy, and David Boud. 1989. "Student Self-Assessment in Higher Education: A Meta-Analysis." *Review of Educational Research* 59 (4): 395–430.

Fink, Dee L. 2003. *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. San Francisco: John Wiley & Sons. Inc.

Flavell, John. 1976. "Metacognitive Aspects of Problem Solving." In *The Nature of Intelligence* (231–36). Hillsdale, NJ: Erlbaum.

Lindsey, Beth A., and Megan Nagel. 2013. "Assessing Students' Metacognitive Calibration with Knowledge Surveys." In *AIP Conference Proceedings* 1513 (1): 258–61). American Institute of Physics. https://doi.org/10.1063/1.4789701.

Mahlberg, Jamie. 2015. "Formative Self-Assessment College Classes Improves Self-Regulation and Retention in First/Second Year Community College Students." *Community College Journal of Research and Practice* 39 (8): 772–83. https://doi.org/10.1080/10668926.2014.922134.

McDevitt, Andrew L., Manisha V. Patel, Brad Rose, and Aaron M. Ellison. 2016. "Insights Into Student Gains From Undergraduate Research Using Pre- and Post-Assessments." *BioScience* 66 (12): 1070–78. https://doi.org/10.1093/biosci/biw141.

Nuhfer, Edward. 1996. "The Place of Formative Evaluations in Assessment and Ways to Reap Their Benefits." *Journal of Geoscience Education* 44: 385–94. https://doi.org/10.5408/1089-9995-44.4.385.

Nuhfer, Edward, Christopher Cogan, Steven Fleischer, Eric Gaze, and Karl Wirth. 2016a. "Random Number Simulations Reveal How Random Noise Affects the Measurements and Graphical Portrayals of Self-Assessed Competency." *Numeracy* 9 (1). https://doi.org/10.5038/1936-4660.9.1.4.

Nuhfer, Edward B., Christopher B. Cogan, Carl Kloock, Gregory G. Wood, Anya Goodman, Natalie Z. Delgado, and Christopher W. Wheeler. 2016b. "Using a Concept Inventory to Assess the Reasoning Component of Citizen-Level Science Literacy: Results From a 17,000-Student Study." *Journal of Microbiology and Biology Education* 17 (1): 143–55. https://doi.org/10.1128/jmbe.v17i1.1036.

Nuhfer, Edward, Steven Fleischer, Christopher Cogan, Karl Wirth, and Eric Gaze. 2017. "How Random Noise and a Graphical Convention Subverted Behavioral Scientists' Explanations of Self-assessment Data: Numeracy Underlies Better Alternatives." *Numeracy* 10 (1). https://doi.org/10.5038/1936-4660.10.1.4.

Nuhfer, Edward, and Delores Knipp. 2003. "4: The Knowledge Survey: A Tool for All Reasons." *To Improve the Academy* 21: 20210331. https://doi.org/10.3998/tia.17063888.0021.006.

Nuhfer, Edward B., and Delores Knipp. 2006. "Re: The Use of a Knowledge Survey as an Indicator of Student Learning in an Introductory Biology Course." *CBE—Life Sciences Education* 5 (4): 313–14. https://doi.org/10.1187/cbe.06-05-0166.

Pintrich, Paul R. 2000. "The Role of Goal Orientation in Self-Regulated Learning." In *Handbook of Self-Regulation*, edited by Monique Boekaerts, Paul R. Pintrich, and Moshe Zeidner, 451–502. https://doi.org/10.1016/b978-012109890-2/50043-3.

Pintrich, Paul R. 2002. "The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing." *Theory Into Practice* 41 (4): 219–25. https://doi.org/10.1207/s15430421tip4104_3.

Ross, John A. 2006. "The Reliability, Validity, and Utility of Self-Assessment." *Practical Assessment, Research, and Evaluation* 11 (10): 1–13. https://doi.org/10.7275/9wph-vv65.

Rushton, Alison. 2005. "Formative Assessment: A Key to Deep Learning?" *Medical Teacher* 27 (6): 509–13. https://doi.org/10.1080/01421590500129159.

Scharff, Lauren F., Chandler-Garcia, Lynne M., Culver, Christopher, Frank, Timothy E., Post, Martiqua, Tetley, Julie, Sloan, Joel A. In Press. "Knowledge Surveys: Developing Self-Assessment, Integrating Effective Learning Strategies, and Enhancing Transparency to Support Student Success." Book chapter accepted for

TILTed Pedagogy: A Collection of Evidence-Based Assignments to Inspire Learning, edited by Shannon Sipes, Mandy Frake-Mistak, Jennifer Friberg. Elon's Center for Engaged Learning (CEL) Open Access Book Series.

Sloan, Joel A., Melissa M. Beauregard, and M. Mark Russell. 2023. "Chapter 4: High Impact Practices in the Civil Engineering Major at the United States Air Force Academy." In *High Impact Practices in Higher Education: International Perspectives (Innovations in Higher Education Teaching and Learning*, edited by Enakshi Sengupta. Emerald Publishing Ltd. https://doi.org/10.1108/S2055-364120230000051004.

Sloan, Joel A., Karin L. Becker, and Timothy E. Frank. 2022. "Implementing a Technical Communication Block in a Civil Engineering Computer Applications Course: Lessons Learned." Proceedings of the 2022 ASEE Annual Conference, Minneapolis, MN. https://peer.asee.org/41191.

Sloan, Joel A., and Timothy E. Frank. 2023. "Student Self-Assessment Data: An Instrument for Program Assessment." *Intersection: A Journal at the Intersection of Assessment and Learning* 4 (2). http://dx.doi.org/10.61669/001c.84511.

Sloan, Joel A., and Timothy E. Frank. 2024. *Relationships Between Student Self-Assessment Ability and Performance*. Paper presented at 2024 ASEE Annual Conference & Exposition, Portland, Oregon. https://peer.asee.org/47937.

Sloan, Joel A., and Lauren F. Scharff. 2022. "Student Self-Assessment: Relationships Between Accuracy, Engagement, Perceived Value, and Performance." *Journal of Civil Engineering Education* 148 (3). https://doi.org/10.1061/(asce)ei.2643-9115.0000063.

Sullivan, Keith, and Cedric Hall. 1997. "Introducing Students to Self-assessment." *Assessment & Evaluation in Higher Education* 22 (3): 289–305. https://doi.org/10.1080/0260293970220303.

Tanner, Kimberly D. 2012. "Promoting Student Metacognition." *CBE—Life Sciences Education* 11 (2): 113–20. https://doi.org/10.1187/cbe.12-03-0033.

Taras, Maddalena. 2002. "Using Assessment for Learning and Learning From Assessment." *Assessment & Evaluation in Higher Education* 27 (6): 501–10. https://doi.org/10.1080/0260293022000020273.

Taras, Maddalena. 2010. "Student Self-Assessment: Processes and Consequences." *Teaching in Higher Education* 15 (2): 199–209. https://doi.org/10.1080/13562511003620027.

Wirth, Karl R., Edward B. Nuhfer, Rachel Watson, Steven C. Fleisher, and Roxanne Bailey. 2021. "Knowledge Surveys: Supporting Students Along Pathways to Self-Directed Learning With Self-Assessment." In *Self Directed Learning: An Imperative for Education in a Complex Society, edited by Elsa Metnz, Dorothy Laubscher, and Jako Olivier*, 127–64. https://doi.org/10.4102/aosis.2021.bk279.06.

Wirth, Karl R., and Dexter Perkins. 2005. "Knowledge Surveys: An Indispensable Course Design and Assessment Tool." *Innovations in the Scholarship of Teaching and Learning*.

Wirth, Karl R., and Dexter Perkins. 2013, October 27. "Learning to Learn." Available from: https://www.montana.edu/rmaher/barrier_courses/Learning%20to%20Learn%20Wirth.pdf.

Zell, Ethan, and Zlatan Krizan. 2014. "Do People Have Insight Into Their Abilities? A Metasynthesis." *Perspectives on Psychological Science* 9 (2): 111–25. https://doi.org/10.1177/1745691613518075.