

RESEARCH

Open Access



Development of the student course cognitive engagement instrument (SCCEI) for college engineering courses

Allyson Barlow^{1*}, Shane Brown², Benjamin Lutz³, Nicole Pitterson⁴, Nathaniel Hunsu⁵ and Olusola Adesope⁶

Abstract

Background: Evidence shows that students who are actively engaged with learning materials demonstrate greater learning gains than those who are passively engaged. Indeed, cognitive engagement is often cited as a critical component of an educational experience. However, understanding how and in what ways cognitive engagement occurs remains a challenge for engineering educators. In particular, there exists a need to measure and evaluate engagement in ways that provide information for instructors to deploy concrete, actionable steps to foster students' cognitive engagement. The present study reports the development and gathering of validation evidence for a quantitative instrument to measure students' in-class cognitive engagement. The instrument was informed by Wylie and Chi's ICAP (Interactive Constructive Active Passive) model of active learning, as well as contextual concerns within engineering courses.

Results: The process followed the classical measurement model of scale development. We provide a detailed overview of the item development and scale validation processes, focusing on the creation of individual subscales to measure different modes of cognition within learning contexts. Multiple rounds of testing the student course cognitive engagement instrument (SCCEI) in college engineering courses provided evidence of validity. This indicated the reliable measurement of student cognitive engagement in the context of notetaking, processing material, and interacting with peers in the classroom. Results suggest differentiating modes of cognitive engagement is indeed applicable when considering students' in-class notetaking and processing of material.

Conclusions: Findings point towards the need for additional engagement scales that expand the instrument's ability to distinguish between particular activities within a mode of engagement as defined by ICAP. The present study contributes to the growing body of literature on cognitive engagement of engineering students. Results address the development of measurement tools with evidence of validity for use in STEM education.

Keywords: Student cognitive engagement, Instrument development, Self-report measurement, Exploratory factor analysis

* Correspondence: allysonbarlow@unr.edu

¹School of Chemical and Materials Engineering, University of Nevada Reno,
1664 N. Virginia St. M, Reno 89557, NV, USA

Full list of author information is available at the end of the article

Introduction

Engineering education research emphasizes the importance of engagement for student learning and academic success (Christenson, Reschly, & Wylie, 2012; Fredricks, Blumenfeld, & Paris, 2004; Jones, Parette, Hein, & Knott, 2010). Researchers prompt educators to innovate and generate engaging courses for the betterment of their students (Chen, Lattuca, & Hamilton, 2008; Chi & Wylie, 2014). Yet, educators are often left to make meaning of what student engagement might look like in their course without the support of the researchers who study engagement—to do so requires educators to interpret theoretical definitions of engagement and apply them to their unique course contexts. One strategy to promote innovation related to engagement in the classroom is to provide educators with tools to measure their success in facilitating student engagement. Tools to measure engagement place the responsibility of theory interpretation on the tool developers, thereby limiting the need for educators to do this themselves. This mitigates concern that educators who inappropriately assimilate theory-based practices into their courses may conclude they are simply ineffective (Henderson & Dancy, 2011).

One foundational component of measurement tools is clearly defining the phenomenon to be measured (DeVellis, 2017). However, a notable challenge is that there are many different, and equally valid, ways of defining and discussing student engagement in extant engagement literature. Craik and Lockhart (1972) discussed engagement in terms of depth of cognitive processing (e.g., *shallow* versus *deep*); recently, theorists have examined engagement in terms of its multifaceted nature (e.g., *behavioral, emotional, and cognitive engagement*) (Appleton, Christenson, Kim, & Reschly, 2006; Fredricks et al., 2004). While these different ways of conceptualizing engagement have been informative in different contexts, determining the most authentic way to assess indicators of engagement that have the most direct and observable bearing on teaching and instruction in the classroom remains a challenge for engineering educators. For example, despite the fact that research has repeatedly indicated the existence of a strong positive relationship between student learning and cognitive engagement, it has been difficult to measure cognitive engagement in the classroom satisfactorily (Chi & Wylie, 2014). This is perhaps the case because a definition for the concept of cognitive engagement has been particularly elusive.

Recently, Chi and colleagues proposed the interactive-constructive-active-passive (ICAP) framework, a model for conceptualizing different dimensions of cognitive engagement (Chi & Wylie, 2014). The framework establishes modes of cognitive engagement that can be

observed as overt behaviors in students. The present study sought to develop an instrument to that leveraged the ICAP framework to indicate the mode of cognitive engagement students exhibited in classroom learning contexts of notetaking, processing material, and interacting with peers. This survey is intended to be applicable to educators who want to better assess student cognitive engagement in their engineering classes, especially as they reflect on the impact of instructional innovations intended to enhance student engagement in their classroom. We also hope that the STEM (Science, Technology, Engineering, and Mathematics) education research community would find the instrument a viable tool to assess depths and quality of cognitive engagement in a range of classroom contexts. The present work provides a detailed overview of how the instrument was developed to measure student cognitive engagement.

Relevant literature

This section briefly discusses how engagement has been defined, emphasizing significant literature contributing to the development of student cognitive engagement theories. We also discuss the ICAP framework, and its usefulness in capturing different dimensions of cognitive engagement. We provide an overview of theoretical perspectives of engagement in extant literature that informed how we operationalized cognitive engagement, while a general overview of survey development literature is discussed in the methods section.

Engagement

Researchers have sought to define engagement in a broadly encompassing manner. Within the scholarship of teaching and learning, engagement is generally construed as specific behaviors that students exhibit within a learning environment that indicate the quality of their involvement or investment in the learning process (Pace, 1998). Some researchers have posited that engagement can be thought of as a meta-construct—one that can be broken down into the components of *behavior*, *emotion*, and *cognition* (Fredricks et al., 2004). Behavioral engagement entails involvement in learning and academic tasks, as well as participation in school-related activities. Emotional engagement encompasses students' affective response, or commitments, to activities in their learning environment. Cognitive engagement, the focus of the present work, refers to the level of psychological investment in the learning process exhibited by the learner (Fredricks et al., 2004). More recent works have included a fourth component of student engagement: *agentic* (Reeve, 2013; Reeve & Tseng, 2011; Sinatra, Heddy, & Lombardi, 2015). Agentic engagement can be defined by the agency or constructive contribution students make towards their learning (Reeve & Tseng, 2011). While

agentic engagement is a relatively new construct, it has been shown to be a statistically significant predictor of achievement in its own right (Reeve, 2013). As is noted by Sinatra et al. (2015), dimensions of engagement co-occur; they direct researchers to be aware that when targeting measurement of one construct of engagement, others are undoubtedly contributing to its measurement.

Exploring the impacts of engagement, broadly defined, on student outcomes such as persistence, migration, self-efficacy, and student performance has been useful to the engineering education research community (Freeman et al., 2014; Kashefi, Ismail, & Yusof, 2012; Ohland, Sheppard, Lichtenstein, Chachra, & Layton, 2008; Sun & Rueda, 2012). However, some have argued that student engagement, as broadly defined in many studies, favors measurement of observable behavioral activities that are not necessarily indicative of students' cognitive investment in the learning process. This, perhaps, is due to the fact that students' behavioral activities are the only aspect of the engagement meta-construct that can be directly observed and thereby assessed (Fredricks et al., 2004). Cognitive and emotional engagement are thus considered latent constructs; they cannot be directly measured, and require more intentional approaches that focus on the measurement of related variables to be captured (McCoach, Gable, & Madura, 2013).

Cognitive engagement

Cognitive engagement is conceptualized in the learning and instruction literature as the psychological investment students make towards learning—which ranges from memorization to the use of self-regulatory strategies to facilitate deep understanding (Fredricks et al., 2004). Irrespective of pedagogical strategies, research shows that meaningful learning is predicated on quality cognitive engagement (Guthrie et al., 2004; Smith, Sheppard, Johnson, & Johnson, 2005). In fact, cognitive engagement is at the hallmark of the *Seven Principles of Good Practice in Undergraduate Education* (Chickering & Gamson, 1987). Among other things, Chickering and Gamson's seven principles, which include active learning and contact between students and faculty, emphasize the importance of cognitive engagement to learning. Deep cognitive engagement has been linked directly to achievement (Greene, 2015). To increase cognitive engagement, students must move from *shallow* cognitive processing to *meaningful* cognitive processing (Craik & Lockhart, 1972). Deep cognitive processing allows for the kind of mental connection and knowledge elaboration that fosters higher-level cognitive learning outcomes, while shallow processing perpetuates rote learning most engendered by lack of robust engagement with the learning materials (Christopher, Walker, Greene, & Mansell, 2005).

Measurement of cognitive engagement

Measuring this important construct is not a new venture, as several education researchers have developed a variety of approaches to assessing students' cognitive engagement. Meece, Blumenfeld, and Hoyle (1988) conceptualized cognitive engagement in terms of student goals and their impact on learning, and thus proposed individual cognitive engagement as a function of learning goals. Inspired by Meece et al., Greene and Miller (1996) developed a measure of *meaningful* and *shallow* cognitive engagement based on a student achievement framework, dubbed the Motivation and Strategy Use Survey. They reported empirical data to support the predictive validity evidence of a measure of cognitive engagement based on a goal-achievement theoretical framework. Their study confirmed a relationship between perceived ability and goals student set for their learning (Greene & Miller, 1996), reaffirming the importance of student cognitive engagement. Validation evidence of this instrument was collected from an educational psychology class and items were general and not engineering-course specific. Appleton et al. proposed a measure of cognitive and psychological engagement that is focused on "students' perceived competence, personal goal setting, and interpersonal relationships" (Appleton et al., 2006, p. 431). Their 30-item Student Engagement Instrument (SEI) was developed based on a context-student engagement-outcomes taxonomy derived from a review of engagement-related literature at the time. The SEI was designed to assess the cognitive engagement of middle school and high school students. Similar to the SEI, the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich, 2015) also provided insight into student cognitive engagement, as defined in terms of motivation. While showing evidence of validity for generalized measure of a student's engagement, both instruments have limited usefulness in specialized contexts such as an engineering course. The SEI does not relate to a specific learning context (e.g., a singular classroom), and the MSLQ does not clearly report on modes of cognition (e.g., at what point is engagement *meaningful*?).

Scoping instruments to measure cognitive engagement as it relates to a course is important if such measurement tools are intended to be used in assessing instructional effectiveness or to evaluate pedagogical practices—particularly given that there have been calls within our research community to modify engineering classes to encourage active learning (i.e., Chi, 2009; Prince, 2004). Some work has already been done to address this need, such as Heddy et al.'s work on the development of an instrument to measure cognitive engagement as students undergo conceptual change (Heddy, Taasoobshirazi, Chancey, & Danielson, 2018). The validity evidence gathered in this study indicates

that cognitive engagement can indeed be measured in the context of a student's activities (e.g., movement from misconception to knowledge, interacting with peers, taking notes, or processing newly introduced material). Yet, as stated by Greene in *Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research*, there are limited tools to measure cognitive engagement in uniquely challenging context of the sciences (Greene, 2015). Chi and Wylie's interactive-constructive-active-passive (ICAP) framework for linking cognitive engagement with learning outcomes provides a theoretical model for operationalizing and measuring cognitive engagement in STEM course-based or classroom learning contexts.

Interactive-constructive-active-passive (ICAP) framework

Chi foregrounded her ICAP framework in a comparative review of research that focused on exploring and classifying active learning activities in the classroom (Chi, 2009). The goal of the work was to "provide a framework differentiating [modes of cognitive engagement], in terms of [students'] overt activities and their corresponding cognitive processes" (Chi, 2009, p. 75). According to Chi, the framework was "not meant to be definitive, but only as a starting point to begin thinking about the roles of overt learning activities and their relationship to internal processes" (Chi, 2009, p. 75). Subsequently, Chi and Wylie further developed their theory of cognitive engagement to include four modes: (1) interactive, (2) constructive, (3) active, (4) passive (Chi & Wylie, 2014). The framework, known as ICAP, links observable behavioral actions of students in different learning situations to distinct modes of cognitive engagement. This way, they moved away from the historical ambiguity associated with broad operational definitions of cognitive engagement.

The ICAP framework operationalizes engagement in terms of activities that are applicable to, and observable in, a number of learning environments. They posit that students show deeper psychological investment in learning and are more cognitively engaged as they move from *passive* to *active* to *constructive* to *interactive* engagement. Invariably, each mode of cognitive engagement requires different levels of knowledge-change processes, or means of acquiring and knowing through learning, that result in increasing levels of meaningful learning from passivity to interactivity (Chi & Wylie, 2014). To illustrate this, Chi and Wylie drew distinctions between modes of engagement by indicating distinctive actions and verbs that characterize each level.

According to the ICAP framework, when passively engaged, students are only oriented towards instruction. For example, students may passively *listen* to a lecture, *read* books, or *watch* a video. Students become actively

engaged when they *repeat*, *rehearse*, or *copy* notes. To be considered constructively engaged, a student must take the original material and manipulate it in a meaningful way. Meaning that, constructively engaged students *reflect*, *integrate*, and *self-explain* concepts. Interactive engagement represents the deepest level of cognitive engagement; students may engage in discussions in which they explain their thoughts and positions to one another. Students may *defend*, *ask*, and *debate* as they interactively engage in the learning context (Chi & Wylie, 2014).

One objective of the ICAP framework is to provide instructors with a tool that enables them to assess the mode of cognitive processing of students. This is accomplished by observing students' learning behaviors as they engage with learning tasks in the classroom (Chi & Wylie, 2014). This task of observing and inferring cognition based on students' overt behaviors can prove daunting to instructors based on factors such as their student population size, time and effort required during class, and understanding of the assumptions of the framework. Furthermore, the usefulness of seeking to observe cognition has been questioned by Appleton, who has suggested that making inferences regarding students' cognitive engagement via observation is not as valid as obtaining students' perspectives on their learning experiences (Appleton et al., 2006).

While the ICAP framework allows for inferences on students' mode of cognitive engagement during a classroom learning activity, a survey-based measurement schema to provide educators aggregated feedback of students' perspective on their own cognitive engagement has yet to be developed. Because the ICAP framework currently relies on the interpretations of an observer, it is limited in its scalability to serve as an effective tool for assessing and evaluating students' cognitive engagement. The ICAP framework also focuses on a learning activity as opposed to the experience of an individual learner. These scalability and specificity challenges create a need to develop measures that both solicit individual student perceptions and are grounded in the ICAP framework.

Evaluating the robustness of the ICAP framework

In response to the ICAP framework, some have designed studies to test the comparative efficacy of instructional methods that encourage each mode of cognitive engagement highlighted by the ICAP framework. Some of the early work to test the ICAP hypothesis was conducted by Menekse and colleagues (Menekse, Stump, Krause, & Chi, 2011; Menekse, Stump, Krause, & Chi, 2013). In their studies, they compared the learning gains of students' contexts that either promoted interactive, constructive, active, or passive learning activities. They found that students had higher and deeper conceptual understanding of materials science and engineering concepts when taught using learning activities that fostered

interactive engagement. Additionally, the test scores demonstrated that students had significant incremental learning gains from passive to active to constructive to interactive activities—which affirms the central ICAP hypothesis.

Wang and colleagues collected and coded data from massive open online course (MOOC) discussion forums, using coding schemes based on the ICAP framework (Wang, Yang, Wen, Koedinger, & Rosé, 2015). They aimed to better understand how students engage in online learning environments that often lack teacher and peer social presence (Akcaoglu & Lee, 2016). They observed that students' active and constructive discussion behaviors significantly predicted students' learning gains, consistent with the ICAP hypotheses (Wang et al., 2015).

The associations between overt behaviors and cognitive engagement underscore the predictive validity of the framework and strengthen the case for using it as a conceptual framework for designing a cognitive engagement instrument. Drawing on the ICAP framework, DeMonbrun and colleagues mapped instructional practices to the four modes of cognitive engagement. Then, students were prompted on their *response* to the instructional practice (i.e., value, positivity, participation, distraction, evaluation) (DeMonbrun et al., 2017). While DeMonbrun used the ICAP framework to indicate the *mode* of the classroom students were in, they did not map the engagement of students to ICAP, or specifically study cognitive engagement. Yet, their work serves to validate that ICAP is a reliable indicator for modes of cognitive engagement in measurement scales.

Because of the importance of cognitive engagement in the development of meaningful learning environments for students, we argue that an optimal instrument would leverage the modes of cognitive engagement proposed by ICAP to provide an empirically reliable tool for assessing engagement in classroom learning contexts. How students interact with one another, take notes, and process material, which are behaviors associated with elements of the ICAP framework, are classroom learning contexts relevant to engineering courses and influenceable by educators. These classroom learning contexts thereby provide a foundational starting point for assessing cognitive engagement.

Survey development and exploratory results

In the following sections, we summarize the development of our instrument to measure student cognitive engagement based on the ICAP framework. This instrument is a part of an ongoing program of research aimed at holistically understanding how STEM students engage with their courses both inside and outside the classroom. While our previous work has offered specific details on modifications made to various versions of our

instrument (Lutz et al., 2018), here we explicate the iterative processes of survey development that led to the current version with evidence of validity and reliability.

Our approach follows recommendations by DeVellis in *Scale development: Theory and applications* (DeVellis, 2017) in large part because the work focuses on the development of measurement tools in educational contexts. DeVellis outlines eight main steps in the development of a scale. We describe in sequence how we executed each of these steps in developing the student course cognitive engagement instrument (SCCEI). The overlapping and iterative nature of scale development is particularly evident in steps 2, 3, and 4 (*generating an item pool, determining a format for measure, and expert review*, respectively)—we generated an item pool, determined item formats, and conducted expert reviews in a concurrent series of activities. Therefore, we present steps 2, 3, and 4 in a single section to allow the reader to follow the logic behind the selected items and measurement format. The 18 items tested for the SCCEI validity evidence are presented at the conclusion of step 4. This paper illustrates how we followed recommended practices in instrument development in an effort to provide a transparent description of the process.

Step 1: Determine clearly what it is you want to measure

The first step in scale development is to think “clearly about the construct being measured” (DeVellis, 2017, p. 105). Obvious though it may sound, it is particularly important in determining the operational definition of the *construct* to measure and the theoretical framework to draw from (Benson & Clark, 1982). Step 1 is important in defining how the intended new instrument differs from any other existing instrument. Identifying the appropriate theoretical framework is germane to item specification and development.

As noted earlier, engagement has been broadly defined and discussed at various levels of specificity in the literature. Researchers emphasized the importance of determining the level of specificity when conceptualizing engagement in an effort to develop a tool to measure the construct (Sinatra et al., 2015). We sought to develop an instrument that assesses cognitive engagement through leveraging the strengths of the ICAP framework (Chi & Wylie, 2014). The ICAP framework is premised on empirical data that associates certain observable behavioral characteristics with cognitive engagement and learning gains. The ICAP framework does not directly address cognition. Rather, behavioral responses are used as proxies for students' cognitive engagement. Utilizing the ICAP framework as our foundational definition for cognitive engagement was strategic given that the framework has been positively received and well-cited within

the engineering education research literatures (e.g., Menekse et al., 2013; Wang et al., 2015).

We based our *construct* and item specificity on how the ICAP framework describes behavioral responses that depict the four levels of cognitive engagement. Because the ICAP framework can be applied to a wide array of learning activities, we looked for learning activities ubiquitously present in engineering courses and influenceable by educators. Constructs to measure were the modes of engagement when students were interacting with peers, taking notes, and processing material in a course. While these constructs do not holistically represent learning in a classroom, or the ICAP framework, they provide an intentional starting point from which to understand modes of cognitive engagement in engineering classes.

After an extensive literature search, a decision was made to allow students to reflect on their own cognition, not only for benefits gained from self-reflection (Nicol & MacFarlane-Dick, 2006), but because their perspective on their own engagement is valuable. Appleton et al. argue that self-report is more useful than observation and rating scales when analyzing emotional and cognitive engagement specifically (Appleton et al., 2006). They argue that observation and teacher-rating scales are unreliable measures of emotional and cognitive engagement due to their highly inferential nature. While self-report may not be reflective of an absolute reality, we are not seeking to prove that students are a perfect judge of their own behaviors (or cognition for that matter). Rather, students' own beliefs about their engagement shape their reality and, in turn, the reality of those seeking to educate them. Therefore, we employed self-report in this study to enrich our understanding of student engagement.

Steps 2, 3, and 4: Generate an item pool, determine the format for measurement, and expert review

In step 2 of DeVellis's model, the developer "generate[s] a large pool of items that are candidates for eventual inclusion in the scale" (p. 109). It is important to generate items that reflect the survey's purpose in sufficient quantity and redundancy. In step 3, we determine the format for measurement, addressing the significance of the type of scaling, the number of response categories, and the time frames associated with the item. Steps 2 and 3 should occur simultaneously, ensuring that items are matched with an appropriate format for measurement. The purpose of step 4 is threefold: (1) have experts rate how relevant the items are to the construct being measured, (2) evaluate for clarity and conciseness, and (3) point out phenomena not addressed in the survey (p. 118).

Here, we present the initial items we developed and their coinciding format for measurement, followed by an overview of the review and modifications made to both,

and finally a presentation of the items and format for measure to study validity in subsequent steps.

Items and measurement schema were developed by our research team of experts in different disciplines including engineering education, psychometrics, educational research, social networking, and faculty change. We conducted virtual meetings monthly for a year as part of ongoing development (steps 2 and 3) and review (step 4) processes. We reached out to Ruth Wylie, a co-author of the published classic work on the ICAP framework, to provide expert review of our items. An additional, substantial piece of the step 4 expert review was interviewing faculty and students. While extensive findings of the feedback generated by both faculty and students can be found elsewhere (Barlow, Lutz, Perova-Mello, Fisher, & Brown, 2018; A. J. Ironside et al., 2017; Ironside, Lutz, & Brown, 2018), here we present the findings most directly related to modifications made to our instrument.

When determining how to specify items for each construct, we considered how ICAP's action verbs were related to interacting with peers, taking notes, and processing material. Consequently, we ensured each item reflects at its root action verbs associated with each of the four levels of cognitive engagement, thus aligning with ICAP.

First, our research team paired the verbs used by Chi and Wylie (2014) with a large range of potential actions or cognitive states (e.g., we paired the generative verbs *compare* and *contrast* with *lecture concepts and course content* to construct items that capture the presence of constructive engagement). Secondly, we generated multiple items to capture each construct being measured. We selected adjectives to minimize overlap between discrete items. Third, in accordance with Benson and Clark (1982), we created a redundant set of items about twice as large as would be needed to capture all the dimensions of the construct we intend to measure. This recommendation is intended to ensure sufficient number of items are retained, as some (poor items) may be lost to the validation process. Lastly, we narrowed our initial pool of items while ensuring the generalizability of those items and their ability to measure each level of cognitive engagement that ICAP prescribes. We ended up with 38 items (see Table 1) to measure four levels of cognitive engagement. Fig. 1 below offers a visual depiction of how the ICAP theory was translated into a redundant set of items using the framework's original verbiage.

DeVellis recommends proactively choosing a time-frame that reflects the objective of the survey. Time-frame highlights the temporal feature of the construct being assessed. Some scales may have a universal time perspective (e.g., stable traits such as locus of control), while others require transient time perspective (e.g., a

Table 1 The 38 items developed in step 2, with targeted mode of ICAP (Continued)

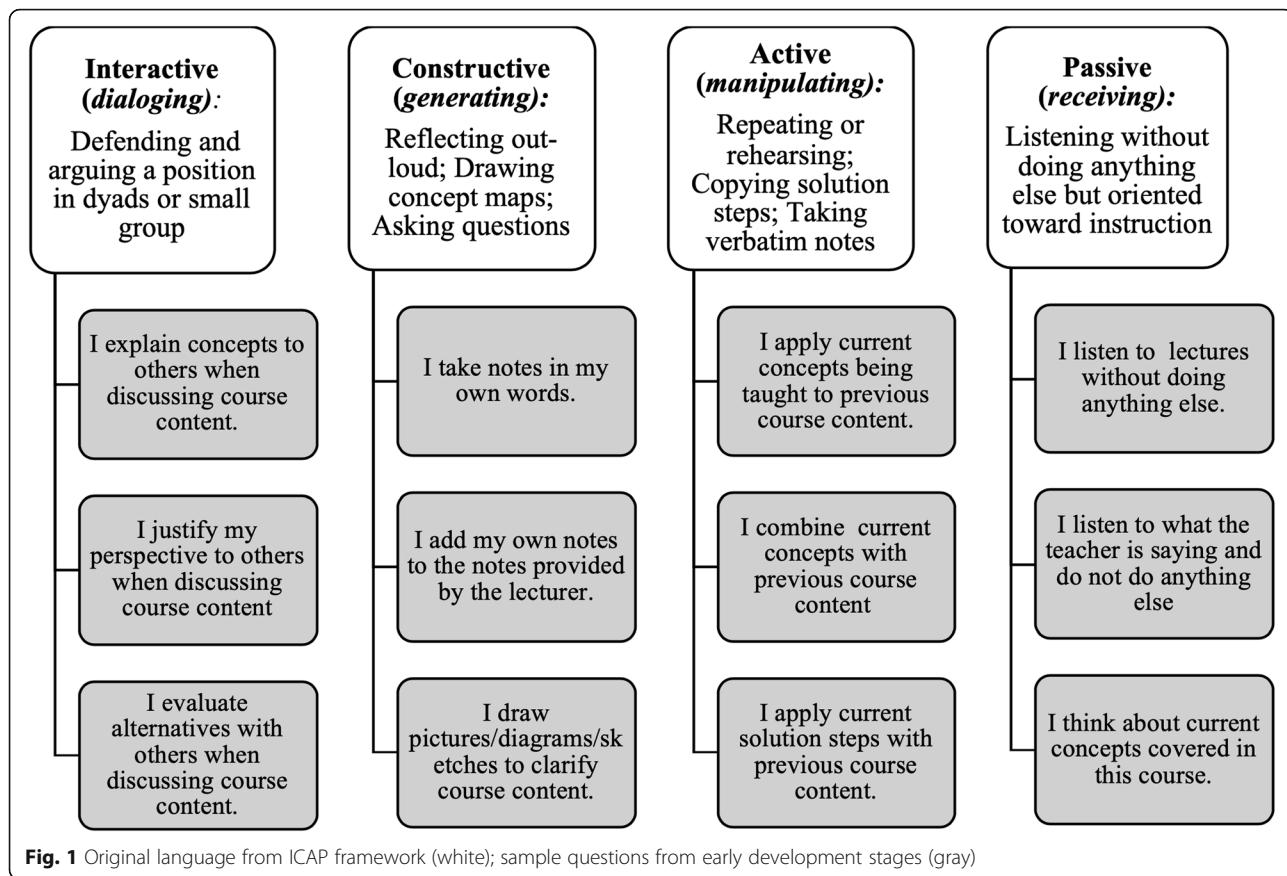
ICAP category	Question
Interactive	I defend my approach to others when discussing course content.
	I discuss my position with others regarding the course content.
	I ask questions to understand other students' perspectives when discussing course content.
	I answer questions describing my perspective when discussing course content
	I explain concepts to others when discussing course content.
	I justify my perspective to others when discussing course content
	I evaluate alternatives with others when discussing course content.
	I do not discuss course concepts with other students.
	I work with other students to understand ideas or concepts regarding course content.
	I take notes in my own words.
Constructive	I add my own notes to the notes provided by the lecturer.
	I draw pictures/diagrams/sketches to clarify course content.
	My course notes include additional content to what the teacher provided.
	I add my own content to the course notes during lecture.
	I consider how multiple ideas or concepts relate.
	I consider how lecture content relates to content from other courses.
	I consider how lecture content relates to course assignments.
	I compare and contrast lecture concepts to concepts from other courses.
	I do not consider how course content relates to other courses.
	I take verbatim notes.
Active	I copy solution steps verbatim.
	I only copy the notes the instructor writes down.
	I do not add my own notes to the course notes.
	I apply current concepts being taught to previous course content.
	I combine current concepts with previous course content
	I apply current solution steps with previous course content.
	I combine current solution steps with previous course content.
	I do not apply course content to previous

Table 1 The 38 items developed in step 2, with targeted mode of ICAP (Continued)

ICAP category	Question
Passive	content.
	I listen to lectures without doing anything else.
	I listen to what the teacher is saying and do not do anything else
	I do not think about course content during lecture.
	I focus my attention on things other than course content during lecture.
	I do not write notes during lecture
	I do not pay attention to course content during lecture.
	I think about current concepts covered in this course.
	I think about previous concepts covered in the course.
	I think about solution steps during the lecture.
	I think about previous solution steps during the lecture.

depression or some activity scale). In determining an appropriate response scale for the survey, we considered possible timeframes that our items could address: a singular incidence/activity, an individual class period, or the aggregate experience of the course. We decided that the Likert scale would address the aggregate experience of the in-class activity of a student within a particular course.

We simultaneously sought to determine an appropriate scale format for the items generated during step 2. Because we were interested in assessing how well respondents believed the items described their learning behaviors, we chose Likert-scale type using the appropriate language "...descriptive of my..." as our response type (Table 2 shows the Likert scale option format that we adopted). This wording was based on a previously developed Likert scale used in educational research related to classroom practices (Walter, Henderson, Beach, & Williams, 2016). In order to determine a convenient level of response options, DeVellis suggests that one considers that respondents are able to *discriminate meaningfully* between response options (DeVellis, 2017, p. 123). For example, it is more conceptually convenient and potentially meaningful to use a 3- or 5-scale Likert type, than to use an 11 or 100-scale type; the fewer the options, the fewer the labels needed (e.g., strongly disagree ... strongly agree) to describe intermediate options. Evidence suggests that different scale lengths have varied benefits and drawbacks (Preston & Colman, 2000; Weijters, Cabooter, & Schilstra, 2010), and there is not overwhelming evidence for exclusive use of a particular scale length.



Initially, we leveraged a 5-point Likert scale to mimic that which had been previously validated; this was later modified to multiple 3-point Likert scales in the dynamic development of the SCCEI

This instrument was built in Qualtrics (2005), an online platform for survey distribution and analytic tools. For initial testing, there were 480 total student responses from 24 different courses. Instructors from the 24 courses were solicited for feedback in a series of interviews. In these interviews, they were asked to share their beliefs on the functionality of the instrument in their classroom environments and what they hoped to gain from the use of such an instrument. Of the 480 students who participated in the study, 13 students volunteered to be interviewed to discuss and justify their responses to survey items.

We learned from these interviews that students often used both in-class and out-of-class justification for their responses to items, all of which were explicitly intended

to relate to their in-class activity. We therefore determined in future iterations of the survey both in-class and out-of-class engagement should be measured simultaneously to explicate the location of the engagement. While we did not seek to validate the out-of-class engagement scale in this study, it provided students with an opportunity to differentiate between where engagement activities took place. Evidence that students did distinguish between the two scales is presented in our other work (A. J. Barlow & Brown, 2019).

Students aggregating their in-class and out-of-class engagement when responding to items resulted in us modifying the format to measure to explicate the location of engagement. We chose to modify the single Likert scale to multiple scales representing different timescales: frequency and duration of activities in-class, and frequency of activities out-of-class. Responses to the in-class frequency scale were the focus of the validation study, while the duration and out-of-class scales are to

Table 2 The 5-point Likert scale associated with the 38 items developed in step 2

Prompt: The following items refer to activities you engage in during class without being directed to do so by your teacher

Not at all descriptive of my in-class activity	Minimally descriptive of my in-class activity	Somewhat descriptive of my in-class activity	Mostly descriptive of my in-class activity	Very descriptive of my in-class activity
--	---	--	--	--

be utilized for future scale development. Three, 3-point Likert scales were used to capture participants' response on multiple timescales (see Table 3).

Subsequently, we conducted a factor analysis to extract an optimum number of factors that underlies the scale and to document validity evidences for the scale. Using an exploratory factor analysis (EFA) ($N = 480$) with principle axis factoring and oblique rotation, tentatively six factors were extracted. Kaiser-Meyer-Olkin (KMO) was 0.932, indicating that the measure of sampling adequacy was sufficient for the EFA. Bartlett's test of sphericity of $\chi^2(703) = 9196.892$, $p < 0.001$, suggesting that there were patterned relationships among the survey items. This analysis was primarily exploratory; decisions to retain items were based in both our conceptualization of engagement and statistical evidence of validity (e.g., eigenvalues greater than 1). We note the ways in which our conceptualization of cognitive engagement (i.e., ICAP) can be applied to the exploratory factors in the following ways.

Items generated appeared to measure a mode of engagement that falls beyond the ICAP framework—disengagement. Chi and Wylie's passive engagement indicates orientation towards instruction, and some students will indeed fall below this threshold. This means some students will not be oriented towards instruction (passively engaged) but will be disengaged with the material altogether. Although we designed some items to be reverse scored, the negatively worded items coalesced around a common factor related to disengagement. We note that some suggest against the practice of including reverse coded/negatively worded items (Roszkowski & Soven, 2010). Therefore, these items related to disengagement were removed. The study of students who fail to engage entirely—those who are disengaged—is beyond the scope of the modes of cognition measured by SCCEI.

Originally, we developed items to measure modes of cognition associated with both notetaking and processing. The preliminary EFA suggested that items factored in alignment with their learning activity (i.e., processing), not simply their mode of cognitive engagement. We noted from preliminary interviews that students seemed to fail to distinguish between various verbs associated with higher-order processing of material, making constructive processing difficult to measure. Beyond this, researcher expertise suggested that notetaking is

inherently *active* in nature; therefore, passive notetaking was not a conceptually reasonable construct for measure by the SCCEI.

We concluded that our items preliminarily measured five distinct phenomena of cognitive engagement: Interactivity with peers, constructive notetaking, active notetaking, active processing, and passive processing. Items were intended to represent an aspect (or indicator) of the mode of cognitive engagement in a classroom with respect to a specific learning experience; the SCCEI may indicate whether a student constructively or actively took notes, actively or passively processed information, and interacted with their peers. Items did not holistically encompass a mode of cognitive engagement, rather they indicated its presence in a given learning activity. Furthermore, the SCCEI does not measure ICAP holistically, rather it relies on ICAP to better understand cognition within classroom notetaking, processing of material, and interactivity with peers.

We leveraged the preliminary EFA to determine the highest correlating items related to our five newly hypothesized phenomena, and iteratively sought evidence for construct and content validity from interview datasets and expert reviews. Items were removed until there were three to four remaining items for each of the five factors. In the end, a set of 18 items remained for validity testing (see Table 4). At the conclusion of steps 2, 3, and 4, we had conducted a thorough review of both our items and format for measure to provide a foundation for seeking validity evidence with subsequent testing.

Step 5: Consider inclusion of validation items

Validation items are intended to limit the influence of factors not related to the construct being measured. A common example of unrelated influence is social desirability, being that certain responses to given items can be seen as more desirable. For example, in the context of our study, validation items might ensure that students are reporting on their engagement, not how their instructor desires for them to be engaged, how peers might view their responses, or how they feel obligated to respond due to social pressures. As we sought to measure cognitive engagement through a distinctive, previously underutilized lens (i.e., ICAP), developing validation items was a substantive task. Therefore, in a follow-up study, we sought external validation of the SCCEI by using it in tandem with a teaching practice

Table 3 Three, 3-point Likert scales; only the in-lecture frequency scale was used for validation

Prompt: I justify my perspective to others when discussing course content.

In lecture (frequency)	In lecture (duration)				Outside the classroom			
Few to no lecture periods	Some lecture periods	Most lecture periods	Little to none of the lecture period	Some of the lecture period	Most of the lecture period	Hardly ever	Some days	Most days

Table 4 18-item version of SCCEI; uses comparison of Likert scales to isolate in-class engagement

ICAP Category	Question
Interactivity with peers	I defend my approach to others when discussing course content.
	I discuss my position with others regarding the course content.
	I explain concepts to others when discussing course content.
	I justify my perspective to others when discussing course content.
Constructive notetaking	I add my own notes to the notes provided by the teacher.
	My course notes include additional content to what the teacher provided.
	I add my own content to the course notes.
Active notetaking	I take verbatim notes (meaning word for word directly from the board/PowerPoint slide/doc camera etc.).
	I copy solution steps verbatim (meaning word for word directly from the board/PowerPoint slide/doc camera etc.).
	I only copy the notes the teacher writes down.
Active processing	I connect current concepts with previous course content.
	I apply current solution steps with previous course content.
	I think about previous concepts covered in the course.
	I consider how multiple ideas or concepts relate.
Passive processing	I pay attention to my teacher or whomever is speaking.
	I follow along with my teacher or whomever is speaking when they discuss examples.
	I listen when my teacher or whomever is speaking.
	I follow along with the activities that take place during the course.

measurement tool. The study consisted of additional data collection beyond that which is present in this study. Results tested for correlations between previously established constructs of teaching best practices and modes of student engagement, adding evidence for the validity of the modes of engagement measured by the SCCEI (Barlow & Brown, n.d.)

Step 6: Administer items to a development sample

In Step 6, the survey was administered to a large sample population. The population ought to be representative of the larger population for which the survey is intended; in the case of our instrument, we sought to develop a survey that could indicate cognitive engagement of students in engineering courses varying in structure, style,

and content. To this end, we recruited 15 engineering courses at eight different institutions that took place at varying points in a four-year curriculum. Institutions ranged in size, emphasis in research, and location. Enrollment in the courses ranged between 33 and 235. As part of reliability testing, an intraclass correlation coefficient (ICC) was generated based on the mean scores of each item for the 15 courses sampled. The ICC estimate and its 95% confidence interval were calculated using SPSS statistical package based on single-rating, consistency (as opposed to absolute agreement), and a 2-way random-effects model. Our ICC value of 0.615 (95% CI, 0.456 to 0.788) indicates that SCCEI items explain approximately 62% of the variation in item scores between courses sampled, and is therefore considered to be moderately reliable (Koo & Li, 2016).

The total population surveyed was 1412 students. After removing responses less than 50% complete, 1004 responses were utilized for analysis, resulting in an overall response rate of 71%. This large sample was randomly split into two groups in order to conduct both an exploratory and confirmatory factor analysis. For a summary of participant demographics, see Table 5 below.

Step 7: Evaluate the items

DeVellis notes that “item evaluation is second perhaps only to item development in its importance” (DeVellis, 2017, p. 139). Evaluation entails examining the manner in which particular items correlate, predict variance, and form reliable scales to measure the desired constructs. To evaluate our items, we conducted factor extraction methods and internal reliability testing in line with recommendations by DeVellis (2017) and Thompson (2004). First, to perform the proper calculations, Likert-type responses were converted into numerical scores. Specifically, we implemented a 3-point scale in which 1 represented low frequency and 3 represented high frequency. Items to which students did not respond were considered null and omitted from subsequent analysis. Exploratory and confirmatory factor analyses were conducted for evidence of validity.

Exploratory factor analysis

We conducted an exploratory factor analysis following recommendations from Thompson (2004) as well as Costello and Osborne (2005) on approximately half of the dataset. The dataset was split into two groups such that the demographics (class size, term sampled, gender, race/ethnicity, etc.) of each set were similar. Although the development of the ICAP framework itself was robust, the structure of, and interaction between, the various modes of engagement is relatively underexplored. Therefore, because we are simultaneously developing a set of scales and operationalizing a theory of student

Table 5 Summary of study population demographics for all student responses

	Study sample	Engr. student population
<i>Institution type</i>		
Large, public northwest R1 institution	47%	-
Moderately-sized, public northwest R2 institution	13%	-
Small, private southwest teaching institution	7%	-
Large, public southwest R1 institution	7%	-
Large, public western teaching institution	7%	-
Large, public southeast R1 institution	7%	-
Small, private northwest teaching institution	7%	-
Moderately-sized, private midwest research institution	7%	-
<i>Course focus</i>		
Civil and construction engineering	40%	8%
General engineering	27%	5%
Chemical engineering	7%	7%
Mechanical engineering	7%	22%
Electrical engineering	7%	17%
Aerospace engineering	7%	4%
Computer science	7%	13%
<i>Academic level</i>		
Freshman	38%	-
Sophomore	9%	-
Junior	23%	-
Senior	30%	-
<i>Gender</i>		
Male	82%	78%
Female	18%	22%
Nonbinary	0.7%	N/A
<i>Race and ethnicity</i>		
Caucasian	78%	67%
Asian	13%	16%
Hispanic/Latinx	4%	12%
Black/African American	4%	4%
Native American	0.4%	0.3%
Pacific Islander	0.2%	0.2%

Note: For *race and ethnicity*, multiple options could be selected. All demographic information collected was optional. Engineering student population data was collected from *Engineering by the numbers* (Yoder, 2018)

engagement, an exploratory factor analysis (EFA) was appropriate. We conducted EFA using SPSS version 24™ with missing values excluded pairwise ($N \sim 495$)—assuming all cases were unique. We utilized principle axis factoring with oblique rotation of items due to the correlation between items. Additionally, we ran reliability tests, using Cronbach's alpha as a metric. Our Kaiser-

Meyer-Olkin (KMO) measure of sampling adequacy was 0.834, indicating a sample sufficient for factor analyses (Cerny & Kaiser, 1977). Bartlett's test of sphericity indicated that the correlation matrix is not an identity matrix and is therefore useful for a factor analysis [$\chi^2(153) = 2794.1, p < 0.001$]. Additionally, with 495 respondents, our ratio of items to respondents was over 20. The number of respondents situates our data within the “very good” sample size as defined by Comrey (1988) and well above the 5 to 10 recommended by Tinsley and Tinsley (1987). We utilized principle axis factoring to better understand the variance between factors (as opposed to principle component analysis, which seeks to better understand how individual items explain overall variance in score).

Several approaches to determining the appropriate number of factors to extract in the dataset have been proposed, the most common being eigenvalue and scree plot (Costello & Osborne, 2005). However, both methods have been criticized as being inadequate to obtaining an optimum number of factors. The parallel analysis (PA) is proposed as a more reliable alternative (Crawford et al., 2010). Parallel analysis represents the amount of variance that would be explained by each factor with completely randomized responses to items; number of responses and items are set equal to the number present in the dataset, and the eigenvalues for each factor are generated. Parallel analysis eigenvalues are compared to the eigenvalues present in the actual dataset—the scree plot. We conducted a parallel analysis based on principle axis factoring (PA-PAF) by running a PA syntax in SPSS® 25 that simulated 5000 parallel data sets from the raw data set using a permutation approach. The PA analysis supports the five-factor model; when factors extracted from the dataset explain more variance than is explained by randomized responses in the parallel analysis, they are considered meaningful. This support is illustrated in the eigenvalue table (Table 6) and the scree plot (Fig. 2) below. We therefore felt confident in extracting five factors. Each factor indicates the presence of a mode of engagement as defined by ICAP in the context of interacting with peers, notetaking, or processing material (see Table 7 for factors extracted).

After determining five factors should be extracted, we looked at the loadings on each factor and the reliability of the items measuring each construct. The absolute value of all factor loadings was above the 0.3 minimum suggested by Hair et al. (Hair, Anderson, Tatham, & Black, 1995). Though the bounds on Cronbach's alpha for reliability are “personal and subjective,” a lower bound of 0.60 is suggested (DeVellis, 2017, p. 145). Reliability for each factor is greater than this 0.60 bound, with some alphas exceeding 0.8. The large number of respondents and small number of items in our scale both influence alpha negatively

Table 6 Results from PA-PAF, which suggest a five-factor model

Root	Raw data	Means	95 th percentile
1	4.251	0.401	0.473
2	1.508	0.330	0.385
3	1.289	0.274	0.321
4	1.066	0.227	0.268
5	0.378	0.184	0.223
6	0.092	0.145	0.180
7	0.052	0.108	0.140
8	0.013	0.073	0.104
9	0.004	0.038	0.066
10	-0.075	0.006	0.033
11	-0.077	-0.027	0.001
12	-0.115	-0.058	-0.033
13	-0.132	-0.090	-0.064
14	-0.159	-0.121	-0.095
15	-0.174	-0.153	-0.128
16	-0.205	-0.187	-0.161
17	-0.230	-0.223	-0.194
18	-0.243	-0.268	-0.234

(lower its value). Our intention was to build a useable instrument, and therefore we traded off a large pool of items for slightly lower reliability. Strong evidence is provided that these five modes of cognitive engagement are indeed distinct and can be defined, at least in part, by differences in behaviors and actions taken to complete in-class activities.

Confirmatory factor analysis

The remaining half of the dataset was used to confirm the findings of the EFA through a confirmatory factor analysis (CFA). Our CFA was conducted in alignment with Brown (2006), who suggests CFAs are useful in verifying both the factors and the relationship of items to factors in questionnaires. We conducted our CFA using AMOS Version 26™ with missing values replaced with means, as CFAs do not allow for missing data ($N = 507$). Our sample of 507 respondents far exceeds the minimum sample suggested by other researchers (Ding, Velicer, & Harlow, 1995; Gorsuch, 1983). We evaluated the model based on the comparative fit index (CFI), Tucker Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR) model fit statistics. A CFI > 0.95 , TLI > 0.95 , RMSEA < 0.06 , and SRMR < 0.08 is recommended for continuous data (Hu & Bentler, 1999).

The CFA yielded the following model fit indices CFI = 0.965, TLI = 0.957, and RMSEA = 0.041, 90% CI [0.033, 0.049], SRMR = 0.0436. Based on the recommendation noted earlier (Hu & Bentler, 1999), the CFA model depicted in Fig. 3 appropriately represents the sample data.

With evidence of good model fit, we analyzed the standardized factor loadings of items and covariances, average variance extracted, and construct reliability of each of the factors. Item reliabilities ranged from 0.531 to 0.884, which exceed the acceptable value of 0.50 (Hair et al., 1995). Covariance between factors ranged between 0.11 to 0.59, falling below the 0.85 threshold that indicates two factors may indeed be measuring a single

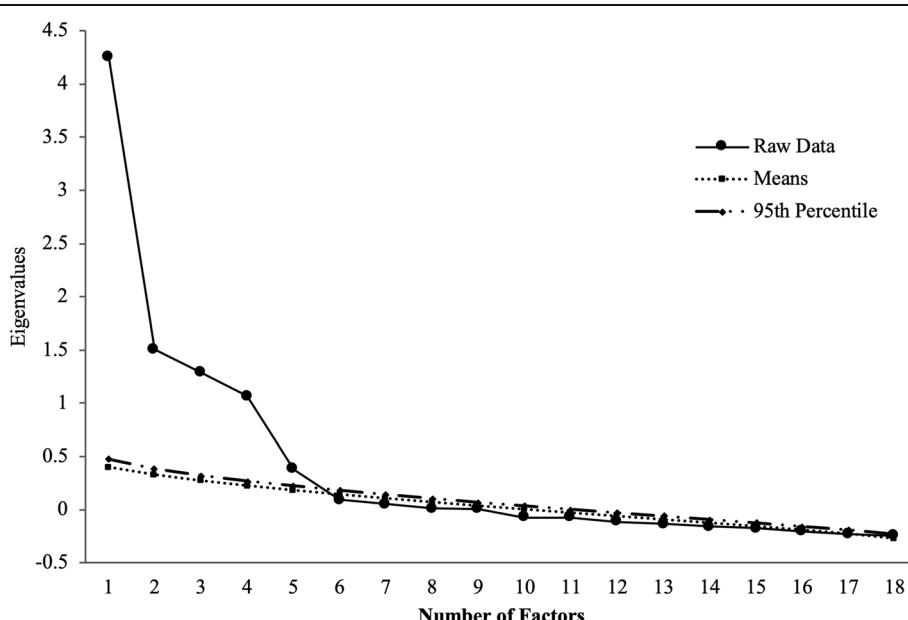
**Fig. 2** Scree plot and results of parallel analysis; points indicate variance extracted for factor(s)

Table 7 Results from the EFA related to frequency of in-class experiences

Instrument item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1: Interactivity with peers					
I defend my approach to others when discussing course content.	0.716				
I discuss my position with others regarding the course content.	0.662				
I justify my perspective to others when discussing course content.	0.656				
I explain concepts to others when discussing course content.	0.564				
Factor 2: Constructive notetaking					
My course notes include additional content to what the teacher provided.	0.835				
I add my own notes to the notes provided by the teacher.	0.769				
I add my own content to the course notes.	0.651				
Factor 3: Active processing					
I think about previous concepts covered in the course.	0.723				
I connect current concepts with previous course content.	0.632				
I apply current solution steps with previous course content.	0.495				
I consider how multiple ideas or concepts relate.	0.42				
Factor 4: Active notetaking					
I take verbatim notes (meaning word for word directly from the board/PowerPoint slide/doc camera etc.)	0.755				
I copy solution steps verbatim (meaning word for word directly from the board/PowerPoint slide/doc camera etc.)	0.675				
I only copy the notes the teacher writes down.	0.567				
Factor 5: Passive processing					
I listen when my teacher or whomever is speaking.	–0.915				
I pay attention to my teacher or whomever is speaking.	–0.796				
I follow along with my	–0.619				

Table 7 Results from the EFA related to frequency of in-class experiences (Continued)

Instrument item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
teacher or whomever is speaking when they discuss examples.					
I follow along with the activities that take place during the course.					–0.543
Eigenvalue	1.677	1.803	4.866	2.063	1.008
Percent variance	9.315	10.017	27.035	11.461	5.601
Cronbach alpha reliability	0.756	0.811	0.741	0.702	0.818

construct (Kline, 2011). Visual representation of the CFA model along with standardized factor loadings and factor covariances can be seen in Fig. 3. The construct reliabilities (CR) ranged from 0.681 to 0.861, and fell above the 0.60 threshold suggested by Bagozzi and Yi (1988). Values for average variance extracted (AVE) fell between 0.421 and 0.610; while it is commonly suggested that AVE values are above 0.5 (Hair, Black, Babin, & Anderson, 2010), if CR values all remain above 0.6, some have suggested items within the factor should be retained (Fornell & Larcker, 1981). Table 8 provides a summary of these results. The cumulative effect of the CFA results added validity evidence to the model suggested by the EFA; we propose that the SCCEI measures five factors of cognitive engaging, indicating the presence of modes of cognition as defined by ICAP with respect to a given classroom experience.

Step 8: Optimize scale length

The length of a scale should be modified in order to increase the reliability of the instrument in step 8. To do so, developers must balance the benefit of a larger number of items increasing reliability with fewer number of items minimizing the burden on the participant (DeVellis, 2017). As we have established, a primary purpose of developing our survey was to provide educators with data on the engagement of students enrolled in their course. Through expert review, we determined that a shortened instrument would be crucial for its practical use. Educators mentioned the importance of an instrument that a majority of students would respond to, requiring that instrument to be of minimal effort to the student. To such an end, we focused our efforts on minimizing the number of items while still maintaining adequate reliability. The Cronbach's alpha of 0.60 and greater for each of the five factors indicated that our scale was both reliable and at a minimum length, and removing any additional items would serve to reduce the overall reliability of the instrument.

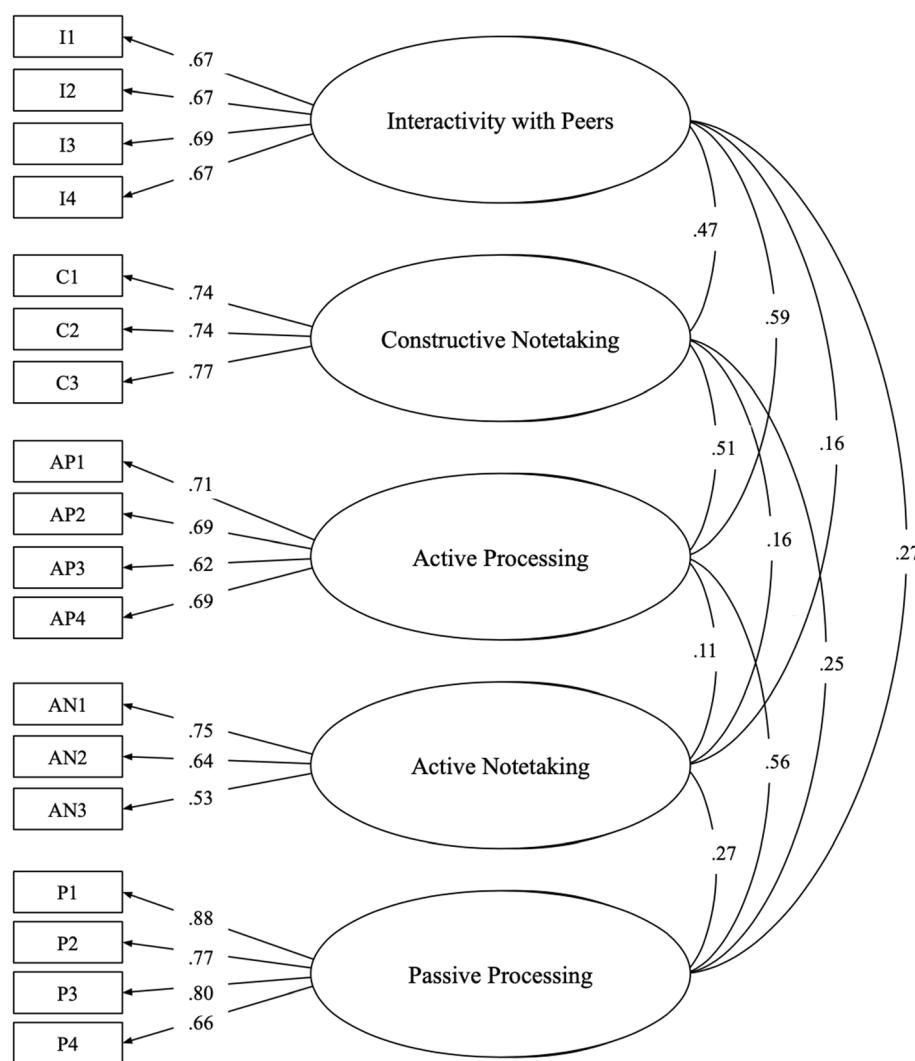


Fig. 3 CFA model; arrows represent standardized factor loadings, factor connectors represent collinearity between factors

Discussion and conclusions

In this study, we reported the development of a new instrument designed to measure the cognitive engagement of engineering students. Chi and Wylie's review of cognitive engagement in the classroom provides a theoretical framework (ICAP) for the development of a new measure of cognitive engagement based on empirically supported relationships between students' learning behaviors, their related cognitive activities, and learning outcomes. A secondary goal of this study was to illustrate the basic steps of scale development as recommended by DeVellis in ways that might serve as a resource for future developers.

Since its publication, the ICAP framework has been favorably received within the engineering education research community. In fact, DeMonbrun and colleagues have initiated a measure of cognitive engagement based on the framework (DeMonbrun et al.,

2017). They posit that a student will identify a classroom as interactive, constructive, active, or passive and respond to that classroom environment cognitively, emotionally, and behaviorally (DeMonbrun et al., 2017). Here, we provide an instrument with validity evidence to determine how students report their own cognitive engagement in the classroom (e.g., the SCCEI allows for modes of engagement in a variety of classroom atmospheres).

Our primary objective was to develop and provide evidence of validity for a measure of cognitive engagement with ties to an empirically verifiable framework (which ICAP provided), and that has broader applications to relevant situations. We set out to develop a tool to measure ICAP in the context of students' cognitive engagement while notetaking, processing material, and interacting with peers; this tool was intended to be useful to educators seeking to deepen understanding of

Table 8 Results from the CFA related to frequency of in-class experiences

Instrument item	Standardized factor loading	Construct reliability	Average variance explained
Factor 1: Interactivity with peers		0.770	0.455
I defend my approach to others when discussing course content.	0.672		
I discuss my position with others regarding the course content.	0.666		
I justify my perspective to others when discussing course content.	0.693		
I explain concepts to others when discussing course content.	0.668		
Factor 2: Constructive notetaking		0.794	0.562
My course notes include additional content to what the teacher provided.	0.738		
I add my own notes to the notes provided by the teacher.	0.738		
I add my own content to the course notes.	0.772		
Factor 3: Active processing		0.772	0.459
I think about previous concepts covered in the course.	0.707		
I connect current concepts with previous course content.	0.689		
I apply current solution steps with previous course content.	0.617		
I consider how multiple ideas or concepts relate.	0.692		
Factor 4: Active notetaking		0.681	0.421
I take verbatim notes (meaning word for word directly from the board/PowerPoint slide/doc camera etc.)	0.752		
I copy solution steps verbatim (meaning word for word directly from the board/PowerPoint slide/doc camera etc.)	0.644		
I only copy the notes the teacher writes down.	0.531		
Factor 5: Passive processing		0.861	0.610
I listen when my teacher or whomever is speaking.	0.884		
I pay attention to my teacher or whomever is speaking.	0.767		
I follow along with my teacher or whomever is speaking when they discuss examples.	0.798		
I follow along with the activities that take place during the course.	0.657		

their students' cognitive engagement in these contexts. Through an extensive collection of validation evidence, this study presents an instrument that allows engineering educators to make valid claims about students' cognition in these instances. We were also able to strengthen claims of the construct validity of the instrument based on the exploratory and confirmatory analyses conducted. We realize, however, that instrument validation is an iterative process that requires multiple studies conducted across diverse populations to strengthen the evidence of the validity of an instrument. As such, we intend to follow up on this study by collecting data to test our validity claims among other student populations, as well as to examine the predictive and concurrent validity of the instrument against established measures or indicators of cognitive engagement. While teacher rating, student observation, and learning outcomes may still remain crucial indicators of cognitive engagement, we envision that the SCCEI could provide researchers with a robust approach for measuring

cognitive engagement with classroom experiences, especially if the intent is to evaluate the effectiveness of particular instructional interventions. Additionally, the frequency scale of the SCCEI allows for educators to prompt students to report on their cognitive engagement at different timescales (i.e., daily, weekly, or term basis).

Implications regarding ICAP framework

Chi and Wylie proposed a pragmatic theoretical lens for differentiating student cognitive engagement in a classroom. The intent of ICAP is to provide a hierarchical description of cognitive engagement that begins with passive engagement (characterized by individualistic learning activities) and progresses to include interactive engagement (characterized by interpersonal, collaborative learning activities). Several studies have examined the central premise of the framework that learning outcomes are positively correlated with increasing levels of cognitive engagement.

In the first phase of our item evaluation, we observed that some respondents seemed to differentiate between verbs related to their experiences of within a course; items related to students' notetaking and their processing of material factored separately, even when related to the same mode of cognition. This suggests that researchers may be able to obtain a more valid self-report of how engaged students are by focusing items to emphasize particular course experiences, such as when a student is taking notes. Furthermore, our work suggests that while both notetaking and processing material are activities that are theoretically aligned with active engagement, operationally, they occur at different frequencies within students. For example, students may actively process material more frequently than they actively take notes in a given course. In this way, we see our work supporting the ICAP framework as it has been previously established: Students' engagement will fluctuate as they participate in different classroom activities. This work provides foundational insight as to what might be involved in the development of a scale to measure the frequency at which student engage at a particulate mode when participating in specific classroom activities.

Chi and colleagues' recent work goes on to note that students have difficulty differentiating between active and constructive activities (Chi et al., 2018). We also observed this phenomenon in our work, suggesting that more work is needed to explore what verbiage clearly elicits a particular mode of cognitive engagement. We see it as important for future work to continue to develop scales that assess the presence of a mode of cognitive engagement related to classroom learning activities (e.g., when students are asked to solve a problem), as well as explore the distinction between closely related modes of cognitive engagement (e.g., active and constructive).

Instructor use of the SCCEI

The SCCEI is bound to appeal to some instructors, as it addressed a subject of broad interest to the educational community—student engagement. Yet, in the initial development of the SCCEI, the primary focus was to align with best practices in scale development. Subsequent focus was placed on generating a useful and useable instrument for instructors. Therefore, instructors who wish to use the SCCEI in their classroom at this time will need to interpret results cautiously and conservatively. We confidently suggest that implementation of the SCCEI will provide some insights on the degree to which student are engaged along its factors. We suggest instructors implement the SCCEI in its entirety, or at a minimum, implement all items related to a given factor of interest.

For instructors wishing to score the SCCEI, responses should first be converted to numerical data. A value of

one (1) should be assigned to the lowest frequency (*few to no lecture periods*), up to a value of three (3) assigned to the greatest frequency (*most lecture periods*). Items related to each factor should then be summed. This sum should then be divided by the total possible number of points in the factor. The result should then be multiplied by 100. In the end, instructors will have calculated a *percent alignment* with each factor.

Instructors should use caution when interpreting the calculated values. We, as the developers, do not know what a specific value means. For example, if the instructor calculates 67% alignment with a given factor, on the surface it means that the average student is engaged at a particular mode 67% of the time. However, in this survey and scale development in general, the precise interpretation of the question or the scale by participants is not known. As a result, the values will give a general sense of student engagement for that particular construct, but not an exact value that can be specifically interpreted. An instructor could implement the instrument in multiple classes over time and eventually get a sense of how student engagement differs in different courses and/or offerings of a single course. This could be used to understand the potential impacts of efforts to improve engagement by the instructor. Instructors may also consider if they are actively trying to encourage a particular form of engagement. For example, if they never ask students to work in groups, then they would not expect to have high scores on the interactive scale. In summary, results must be interpreted very carefully and in context to the classroom environment.

Limitations

SCCEI items are based on the theoretical framework of ICAP, yet these items are limited in their ability to holistically represent a mode of cognitive engagement. Importantly, interactivity with peers indicates that the student reports *potential* for interactive engagement; only nuanced observation and/or discussion with the student would allow for insight as to the level with which they were interactively engaged with the material while interacting with peers. Furthermore, these items do not span the extent of learning activities that generate cognitive engagement in the classroom; originally, more items were included to capture a broader range of engagement activities, but difficulty was encountered in generating reliable factors from these items. Future, substantive work may wish to validate additional items to more holistically capture ICAP modes of cognitive engagement.

The measurement scale likewise limited the granularity at which cognitive engagement can be understood from the SCCEI. We envisaged that using three, 5 or more point Likert scales concurrently for each survey

item would overburden respondents and hinder response rate (Preston & Colman, 2000; Weijters et al., 2010). This rational led to our use of a 3-point Likert scale. Future studies make to wish to add additional response categories to the Likert scale and again test for evidence of validity. This may provide more reliable, meaningful responses than the 3-point scale of this study. Despite limitations, the SCCEI provides educators with meaningful insight as to the presence of various modes of cognitive engagement in different classroom learning experiences.

Our sample was comprised entirely of engineering students. Although we designed our instrument to be useful across STEM courses, validation evidence is needed for the use of the SCCEI outside of engineering. Our sample reasonably represented the population demographics, but Caucasian males remained overrepresented compared to national averages in engineering (see Table 5 above). Further work is needed to understand the nuanced ways in which underrepresented groups cognitively engage, particularly as to how it may differ from normative groups. Valid claims can be made insofar as an instrument provides validity evidence; our broad sampling of courses across many institutions allows the SCCEI to make valid claims about engineering students in general, but limits its ability to provide valid claims about specific engineering student populations. Civil and general engineering were overrepresented in our sample, while mechanical and electrical were underrepresented. Additionally, some disciplines were not represented at all and may require future studies on validity. Generalizability of all findings presented is limited by the sample recruited for the study.

Future direction

Validation efforts involve iterative evaluation and improvement of an instrument in order to improve its psychometric soundness. Currently, we have only three or four items to assess each of the five factors. To further improve the reliability of the subscale comprising the instrument, we intend to create and test additional items for each construct on the scale. In this study, we have examined the structural validity of the subscales on our instrument. Subsequently, we intend to conduct other studies to further strengthen the validity evidence of our instrument. We look to establish its construct and predictive validity by examining measures and proxies of cognitive engagement across a broader sample of students. Efforts will also be made, in the future, to determine the ability of the instrument to effectively discriminate cognitively engaged students from those who are not.

Because we intend to expand the items on the instruments to improve the reliability of the sub-scales, future work would consider focusing on a single subscale (i.e., interactivity with peers) and developing items to related

constructs that could be administered apart from the larger scale for specific research needs and minimize the need to administer the entire cognitive engagement when that is not desired. We posit that more studies are needed to better understand the interplay between engagement inside and outside of the classroom on other variables that mediate or moderate student learning and performance in engineering—especially in specialized learning contexts, (e.g., flipped classrooms). More work is needed to develop scales to indicate the presence of ICAP modes of engagement beyond classroom walls.

Our intent was to develop a scale to measure a construct (cognitive engagement) that is subsumed within a meta-construct (student engagement) drawing upon a theoretical framework that has a strong empirical support. We envision that our study will inspire others to create scales based on empirically grounded theory to measure other constructs germane to engineering education research that are subsumed in broader meta-constructs.

Conclusion

The present study seeks to report our effort to validate a new instrument of cognitive engagement with a literature-based theoretical framework. Our scale development efforts were informed by DeVellis' research. As we explored DeVellis' recommendations, we demonstrated that identifying the relevant literature plays a major part in scale planning development. Consequently, we aligned, and reasonably integrated Chi and Wylie's ICAP framework of cognitive engagement (Chi & Wylie, 2014) to situate the purpose and scope of our instrument. This theoretical alignment was important to the authenticity and the validity of our scale. We demonstrated the importance of engaging experts and target stakeholders in increasing the content validity of the scale being developed, and perhaps more in creating an instrument that is relevant and has a broader application. To create a good scale, one must be open to revising items on the scale, and may have to make the decision to remove poorly performing items. Further, items may not be bound together on the same construct even though they were designed to do so. In fact, it is possible that different factor patterns may emerge, which then would require some theoretical framework to calibrate or interpret.

Finally, we intended to develop a four-factor instrument. However, the items comprising the proposed four-factor scale developed emerged into five factors that are theoretically relevant to the overarching framework on which the instrument was conceptualized. Our instrument provides new perspectives on the ICAP framework and extends its application for scalability to broader contexts. We are hopeful that this study will inspire other innovative research and development in the

area of student cognitive engagement, particularly in engineering education.

Abbreviations

ICAP: Interactive, constructive, active, passive; STEM: Science, technology, engineering, and mathematics; SEI: Student Engagement Instrument; MSLQ: Motivated Strategies for Learning Questionnaire; MOOC: Massive open online course; KMO: Kaiser-Meyer-Olkin; SCCEI: Student course cognitive engagement instrument; EFA: Exploratory factor analysis; ICC: Intraclass correlation coefficient; PA: Parallel analysis; PA-PAF: Parallel analysis based on principle axis factoring; CFA: Confirmatory factor analysis; CF: Comparative fit index; TLI: Tucker Lewis index; RMSEA: Root mean square error of approximation; SRMR: Standardized root mean square residual; AVE: Average variance extracted; CR: Construct reliabilities

Acknowledgements

The authors would like to thank all of the participants of the study, including both instructors who were willing to distribute the survey in their courses and students who agreed to complete the SCCEI.

Authors' contributions

AB was the primary author of the work in this manuscript. AB managed recruitment, distribution, analysis, and write-up of the result of this study. SB oversaw the data collection, analysis, and write up of the work. SB managed project personnel, including guiding the team towards the end goal of an instrument with evidence of validity. BL conducted multiple rounds of data collection and analysis. BL initiated the use of the development framework used in this manuscript and provided feedback on manuscript iterations. NP guided the early development of the SCCEI. She conducted several rounds of pilot testing and made revisions to the instrument that proved critical to its final development and validation efforts. NP also reviewed all drafts of this work and provided revisionary feedback. NH and OA provided substantial guidance on the statistical methods used in this work. Additionally, they provided multiple rounds of revisions and reviews of the work. All authors read and approved the final manuscript.

Funding

This project was funded under NSF grant number 1664250. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

All ethics were overseen by the Institutional Review Board; participants were invited to consent to participate and allowed to opt out of participation with no consequence at any time.

Consent for publication

All participants were informed that consent included agreeance to publication of finding in aggregate. No personalized data is shared in publication materials.

Competing interests

The authors declare that they have no competing interests

Author details

¹School of Chemical and Materials Engineering, University of Nevada Reno, 1664 N. Virginia St. M, Reno 89557, NV, USA. ²School of Civil and Construction Engineering, Oregon State University, 101 Kearney Hall, Corvallis 97331, OR, USA. ³California Polytechnic State University, 1 Grand Avenue, Building 13-221, San Luis Obispo 93407, CA, USA. ⁴Department of Engineering Education, 369 Goodwin Hall, 635 Prices Fork Road, Blacksburg 2406, VA, USA. ⁵College of Engineering, University of Georgia, Athens 99164, GA, USA. ⁶Department of Kinesiology and Educational Psychology, College of Education, Washington State University, Room 356, Cleveland Hall, Pullman 99164, WA, USA.

Received: 12 November 2019 Accepted: 15 April 2020
Published online: 19 May 2020

References

- Akcaoglu, M., & Lee, E. (2016). Increasing social presence in online learning through small group discussions. *International Review of Research in Open and Distance Learning*, 17(3), 1–17. <https://doi.org/10.19173/irrod.v17i3.2293>
- Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology*, 44, 427–445. <https://doi.org/10.1016/j.jsp.2006.04.002>
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74–94. <https://doi.org/10.1007/BF02723327>
- Barlow, A., & Brown, S. A. (n.d.). Correlations between modes of student cognitive engagement and instructional practices in undergraduate STEM courses. *In Review*.
- Barlow, A. J., & Brown, S. A. (2019). Work in progress: measuring student cognitive engagement using the ICAP framework in and outside of the classroom WIP: Measuring student cognitive engagement using the ICAP framework in and outside of the classroom. *ASEE Annual Conference and Exposition Proceedings*.
- Barlow, A. J., Lutz, B., Perova-Mello, N., Fisher, K. Q., & Brown, S. (2018). Factors of sensemaking affecting engineering faculty's decision to use the in-class cognitive engagement survey. *International Journal of Engineering Education*, 34(5), 1658–1670.
- Benson, J., & Clark, F. (1982). A guide for instrument development and validation. *The American Journal of Occupational Therapy*, 36, 789–800.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.
- Cerny, B. A., & Kaiser, H. F. (1977). A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate Behavioral Research*. https://doi.org/10.1207/15327906mbr1201_3
- Chen, H. L., Lattuca, L. R., & Hamilton, E. R. (2008). Conceptualizing engagement: Contributions of faculty to student engagement in engineering. *Journal of Engineering Education*, 97(3), 339–353. <https://doi.org/10.1002/j.2168-9830.2008.tb00983.x>
- Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105. <https://doi.org/10.1111/j.1756-8765.2008.01005.x>
- Chi, M. T. H., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., ... Yaghmourian, D. L. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive Science*, 42, 1777–1832. <https://doi.org/10.1111/cogs.12626>
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. <https://doi.org/10.1080/00461520.2014.965823>
- Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 1, 3–7. [https://doi.org/10.1016/0307-4412\(89\)90094-0](https://doi.org/10.1016/0307-4412(89)90094-0)
- Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *Handbook of research on student engagement*. Retrieved from https://www.researchgate.net/profile/Azkananda_Widiasani/publication/310773130_Handbook_of_Student_Engagement/links/5836a0dd08aed45931c772b7/Handbook-of-Student-Engagement.pdf
- Christopher, B., Walker, O., Greene, B. A., & Mansell, R. A. (2005). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learning and Individual Differences*, 16. <https://doi.org/10.1016/j.lindif.2005.06.004>
- Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*. <https://doi.org/10.1037/0022-006X.56.5.754>
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- Crawford, A. V., Green, S. B., Levy, R., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement*. <https://doi.org/10.1177/0013164410379332>

- DeMonbrun, M., Finelli, C. J., Prince, M., Borrego, M., Shekhar, P., Henderson, C., & Waters, C. (2017). Creating an instrument to measure student response to instructional practices. *Journal of Engineering Education*, 106(2), 273–298. Retrieved from <https://doi.org/10.1002/jee.20162>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (Fourth Ed.). Los Angeles: Sage Publications.
- Ding, L., Velicer, W. F., & Harlow, L. L. (1995). Effects of estimation methods, number of indicators per factor, and improper solutions on structural equation modeling fit indices. *Structural Equation Modeling: A Multidisciplinary Journal*. <https://doi.org/10.1080/10705519509540000>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39. <https://doi.org/10.2307/3151312>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Source: Review of Educational Research*, 74(1), 59–109. Retrieved from <http://www.jstor.org/stable/3516061>
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410–8415. <https://doi.org/10.1073/pnas.1319030111>
- Gorsuch, R. L. (1983). Factor analysis. *Work Study*. <https://doi.org/10.1108/EUM00000000002688>
- Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist*, 50(1), 14–30. <https://doi.org/10.1080/00461520.2014.989230>
- Greene, B. A., & Miller, R. B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology*, 21, 181–192.
- Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., ... Tonks, S. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology*, 96(3), 403–423. <https://doi.org/10.1037/0022-0663.96.3.403>
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). Multivariate data analysis: with readings. In *Statistics*. <https://doi.org/10.2307/2983017>
- Hair, J. F., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River: Prentice-Hall, Inc.
- Heddy, B. C., Taasoobshirazi, G., Chancey, J. B., & Danielson, R. W. (2018). Developing and validating a conceptual change cognitive engagement instrument. *Frontiers in Education*, 3(June), 1–9. <https://doi.org/10.3389/feduc.2018.00043>
- Henderson, C., & Dancy, M. H. (2011). Increasing the impact and diffusion of STEM education innovations. *Characterizing the Impact and Diffusion of Engineering Education Innovations Forum*. Retrieved from <http://create4stem.msu.edu/sites/default/files/discussions/attachments/HendersonandDancy10-2010.pdf>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1). <https://doi.org/10.1080/10705519909540118>
- Ironside, A., Lutz, B., & Brown, S. (2018). Student perspectives on cognitive engagement: Preliminary analysis from the course social and cognitive engagement surveys. *ASEE Annual Conference and Exposition Proceedings*.
- Ironside, A. J., Pitterson, N. P., Brown, S. A. P. E., Fisher, K. Q., Gestson, S. L., Simmons P.E., D. R., & Olusola, A. (2017). Incorporating faculty sense making in the implementation and modification of an instrument to measure social and cognitive engagement. *ASEE Annual Conference and Exposition Proceedings*. Retrieved from <https://peer.asee.org/incorporating-faculty-sense-making-in-the-implementation-and-modification-of-an-instrument-to-measure-social-and-cognitive-engagement>
- Jones, B. D., Paretti, M. C., Hein, S. F., & Knott, T. W. (2010). An analysis of motivation constructs with first-year engineering students: Relationships among expectancies, values, achievement, and career plans. *Journal of Engineering Education*, 99(4), 319–336. <https://doi.org/10.1002/j.2168-9830.2010.tb01066.x>
- Kashefi, H., Ismail, Z., & Yusof, Y. M. (2012). Supporting engineering students' thinking and creative problem solving through blended learning. *Procedia - Social and Behavioral Sciences*, 56(Icthe), 117–125. <https://doi.org/10.1016/j.sbspro.2012.09.638>
- Kline, R. B. (2011). *Methodology in the social sciences. Principles and practice of structural equation modeling* (3rd, Ed.). NY, New York: Guilford Press.
- Koo, T. K., & Li, M. Y. (2016). A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Lutz, B. D., Ironside, A. J., Hunsu, N., Green, C. J., Brown, S. A., & Adesope, O. O. (2018). Measuring engineering students' in-class cognitive engagement: Survey development informed by contemporary educational theories. *ASEE Annual Conference and Exposition Proceedings*.
- McCoach, D. , Gable, R. , & Madura, J. . (2013). *Instrument development in the affective domain: School and corporate applications* (3rd ed.). <https://doi.org/10.1017/CBO9781107415324.004>
- Meece, J. L., Blumenfeld, P. C., & Hoyle, R. H. (1988). Students' goal orientations and cognitive engagement in classroom activities. *Journal of Educational Psychology*, 80(4), 514–523. Retrieved from <http://psycnet.apa.org.ezproxy.proxy.library.oregonstate.edu/fulltext/1989-17194-001.pdf>
- Menekse, M., Stump, G., Krause, S. J., & Chi, M. T. H. (2011). Implementation of differentiated active-constructive-interactive in an engineering classroom. *ASEE Annual Conference and Exposition*.
- Menekse, M., Stump, G. S., Krause, S., & Chi, M. T. H. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education*, 102(3), 346–374. <https://doi.org/10.1002/jee.20021>
- Nicol, D., & MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://doi.org/10.1080/03075070600572090>
- Ohland, M. W., Sheppard S. D., Lichtenstein, G., Eris, O., Chachra, D, Layton. (2008). Persistence, engagement, and migration in engineering programs. *Journal of Engineering Education*, (December), 260–278. <https://doi.org/10.1002/j.2168-9830.2008.tb00978.x>
- Pace, R. C. (1998). Recollections and reflections. In J. C. Smart (Ed.), *Higher Education: Handbook of theory and research Vol. XIII* (pp. 1–34). NY: Agathon Press.
- Pintrich, P. R. (2015). Motivated strategies for learning questionnaire (MSLQ). *Mediterranean Journal of Social Sciences*, 6(1), 156–164. <https://doi.org/10.5901/mjss.2015.v6n1p156>
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1–15. [https://doi.org/10.1016/S0001-6918\(99\)00050-5](https://doi.org/10.1016/S0001-6918(99)00050-5)
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223–231. <https://doi.org/10.1002/j.2168-9830.2004.tb00809x>
- Qualtrics. (2005). Qualtrics. Provo, Utah.
- Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology*, 105(3), 579–595. <https://doi.org/10.1037/a0032690>
- Reeve, J., & Tseng, C. M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, 36(4), 257–267. <https://doi.org/10.1016/j.cedpsych.2011.05.002>
- Roszkowski, M. J., & Soven, M. (2010). Shifting gears: Consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment and Evaluation in Higher Education*, 35(1), 117–134. <https://doi.org/10.1080/02602930802618344>
- Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, 50(1), 1–13. <https://doi.org/10.1080/00461520.2014.1002924>
- Smith, K. A., Sheppard, S. D., Johnson, D. W., & Johnson, R. T. (2005). Pedagogies of engagement: Classroom-based practices. *Journal of Engineering Education*, 94(1), 87–101. <https://doi.org/10.1002/j.2168-9830.2005.tb00831x>
- Sun, J. C. Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology*, 43(2), 191–204. <https://doi.org/10.1111/j.1467-8535.2010.01157.x>
- Thompson, B. (2004). Exploratory and confirmatory factor analysis: Understanding concepts and applications. *American Psychological Association*. <https://doi.org/10.1177/0146621606290168>
- Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*. <https://doi.org/10.1037/0022-0202.016.34.414>
- Walter, E. M., Henderson, C. R., Beach, A. L., & Williams, C. T. (2016). Introducing the postsecondary instructional practices survey (PIPS): A concise, interdisciplinary, and easy-to-score survey. *CBE—Life Sciences Education*, Winter 2015(15: ar53). <https://doi.org/10.1187/cbe.15-09-0193>
- Wang, X., Yang, D., Wen, M., Koedinger, K., & Rosé, C. P. (2015). Investigating how student's cognitive behavior in MOOC discussion forums affect learning

- gains. *International Conference on Educational Data Mining*. Retrieved from <https://files.eric.ed.gov/fulltext/ED560568.pdf>
- Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, 27(3), 236–247. <https://doi.org/10.1016/j.ijresmar.2010.02.004>
- Yoder, B. L. (2018). Engineering by the numbers. *American Society for Engineering Education*, 11–47. Retrieved from https://www.asee.org/papers-and-publications/publications/14_11-47.pdf%5Cnpapers2//publication/uuid/CE4279B3-6EAE-424F-9681-50CBB375C444

Publisher's Note

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

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com